

Humboldt-Universität zu Berlin

Institut für Bibliotheks- und Informationswissenschaft

DISSERTATION

Metadatenbasierte Kontextualisierung architektonischer 3D-Modelle

zur Erlangung des akademischen Grades

Doctor philosophiae (Dr. phil.)

eingereicht an der Philosophischen Fakultät I

von Ina Blümel, Dipl.-Ing. M.Sc.

aus Hannover

Dekan: Prof. Michael Seadle, PhD

Gutachter:	1. Prof. Dr. Stefan Gradmann
	2. Prof. Dr. techn. Dieter W. Fellner
eingereicht:	07.01.2013
Datum der Promotion:	11.02.2013

Zusammenfassung

Digitale 3D-Modelle der Architektur – z.B. Modelle von Gebäuden, Inneneinrichtungsgegenständen und Bauteilen – haben innerhalb der letzten fünf Jahrzehnte sowohl die analogen, auf Papier basierenden Zeichnungen als auch die physischen Modelle aus ihrer planungs-, ausführungs- und dokumentationsunterstützenden Rolle verdrängt.

Als Herausforderungen bei der Integration von 3D-Modellen in digitale Bibliotheken und Archive sind zunächst die meist nur rudimentäre Annotation mit Metadaten seitens der Autoren und die nur implizit in den Modellen vorhandenen Informationen zu nennen. Aus diesen Defiziten resultiert ein aktuell starkes Interesse an inhaltsbasierter Erschließung durch vernetzte Nutzergruppen oder durch automatisierte Verfahren, die z.B. aufgrund von Form- oder Strukturmerkmalen eine automatische Kategorisierung von 3D-Modellen anhand gegebener Schemata ermöglichen. Die teilweise automatische Erkennung von objektinhärenter Semantik vergrößert die Menge an diskreten und semantisch unterscheidbaren Einheiten. Darüber hinaus sind digitale 3D-Modelle zumeist hierarchisch aufgebaut; sie enthalten weitere komplexe Modelle, die wiederum in sich geschachtelt sein können und in einzelnen Fällen einen eigenständigen Nachweis als 3D-Modell wünschenswert machen. 3D-Modelle als Content im World Wide Web können sowohl untereinander als auch mit anderen textuellen wie nichttextuellen Objekten verknüpft werden, also Teil von aggregierten Dokumenten sein. Eine weitere Notwendigkeit ist die Vernetzung mit inhaltlich relevanten Ereignissen, Orten, Begriffen, Personen oder realen Objekten sowie die explizite Beschreibung der Relationen zwischen dem Modell selbst und diesen Entitäten seines spezifischen Kontextes.

Die Aggregationen bzw. der Modellkontext sowie die inhärenten Entitäten erfordern Instrumente der Organisation, um dem Benutzer bei der Suche nach Informationen einen Mehrwert zu bieten, insbesondere dann, wenn textbasiert nach Informationen zum Modell und zu dessen Kontext gesucht wird. In der vorliegenden Arbeit wird ein Metadatenmodell zur gezielten Strukturierung von Information

entwickelt, welche aus 3D-Architekturmodellen gewonnen wird. Mittels dieser Strukturierung kann das Modell mit weiterer Information vernetzt werden. Die Anwendung etablierter Ontologien sowie der Einsatz von URIs machen die Informationen nicht nur explizit, sondern beinhalten auch eine semantische Information über die Relation selbst, sodass eine Interoperabilität zu anderen verfügbaren Daten im Sinne der Grundprinzipien des Linked-Data-Ansatzes gewährleistet wird. Diese Herangehensweise hat im Gegensatz zu einem Ansatz, der Metadaten als Records auffasst, das Potenzial, Relationen zu jeglichen modellrelevanten Entitäten im Suchraum herzustellen und zugleich diese Relationen für weitere wissensbildende Prozesse verfügbar zu machen.

Abstract

Digital 3D models from the domain of architecture – e.g. models describing buildings, furniture or building blocks – have replaced analogue paper-based drawings as well as haptic scale models bit by bit during the last five decades.

The main challenges for integrating 3D models in digital libraries and archives are posed by mostly only sparse annotation with metadata provided by the author and the fact that information is only implicitly available. This has recently led to an increased interest in context-based indexing using automatic approaches as well as social tagging. Computer based approaches usually rely on methods from artificial intelligence including machine learning for automated categorization based on geometric and structural properties according to a given classification scheme. The partially automated recognition of model-inherent semantics increases the number of discrete and semantically distinguishable entities. Furthermore, 3D models usually provide an intrinsic hierarchy, i.e. they consist of complex sub-models which might again consist of sub-models on their own. This type of hierarchical structure calls for explicit documentation of the sub-contents contained in individual cases. 3D models as parts of the World Wide Web can be interlinked which each other. Additionally, they can be linked to other textual or non-textual objects and are therefore considered to be parts of aggregated documents. Another need is to establish links to content relevant entities like e.g. events, places, concepts, people, or real objects and the explicit description of the relationships between the model itself and these entities of its specific context.

Aggregations as well as the model context along with inherent entities require efficient tools for organization in order to provide real additional benefits for the user during its quest for information. Especially for text-based search on information about a 3D model and its context, a metadata model is an indispensable tool regarding the above described challenges. In this work we develop a metadata model for specific structuring of information, which is obtained from 3D architectural models. Using this structure, the model can be linked to further information. The application of established ontologies and the use of URIs make the informa-

tion not only explicitly, but also provide semantic information about the relation itself. By that, interoperability according to the principles of the LOD approach is guaranteed. In contrast to representing metadata as a set of records, our approach can potentially establish relationships to any entities in the search space that are relevant to the 3D model. Additionally, these relationships are made available for subsequent knowledge-creating processes.

Schlagwörter:

Metadaten, 3D-Modelle, Kontextualisierung, Linked Data

Keywords:

Metadata, 3D models, Contextualization, Linked Data

Inhalt

Abbildungsverzeichnis	xi
Abkürzungsverzeichnis	xiii
Einleitung und Zielsetzung	1
1 Digitale Bibliotheken	5
1.1 Definition	5
1.2 Herausforderungen	6
1.2.1 Informationsvolumen	6
1.2.2 Erweiterung des Suchraums	7
1.2.3 Medienformate	9
1.2.4 Strukturelle Aspekte / Dokumentenmodelle	10
1.3 Fazit	11
2 3D-Modelle	14
2.1 Definition	14
2.1.1 Original und Modell	14
2.1.2 Anwendungskontexte → Architektur	15
2.2 Digitale Modellierung	17
2.2.1 Planungsphasen und Modelltypen	18
2.2.2 IT-Entwicklung und Aspekte der Computerunterstützung im Planungsprozess	20
2.3 Strukturelle Eigenschaften und technische Attribute	22
2.3.1 Anzahl an Dateien	22
2.3.2 Geometrische Repräsentation	23
2.3.3 3D-Modellierungsprogramme und Formate	26
2.3.4 Modellaufbau	28
2.4 Suche nach 3D-Modellen	29
2.4.1 Suche im Entwurf: Case-based Reasoning	30
2.4.2 Inspiration und Integration	30
2.4.3 Befragungsergebnisse	32
2.5 Anforderungen an Metadaten für 3D-Modelle	36
3 Informationsextraktion und Kontextualisierung	38
3.1 Retrieval	38

3.2	Wissen organisieren	39
3.2.1	Daten – Information – Wissen	39
3.2.2	Erschließung (traditionell)	41
3.2.3	Records und „Datensilos“	42
3.3	Metadaten: Beschreibende Funktion	44
3.3.1	Aktualisierte begriffliche Annäherung	44
3.3.2	Anwendung	46
3.3.3	Struktur / Elementtypen	47
3.3.4	Ebenen	48
3.4	Information extrahieren	50
3.4.1	Wissen in 3D-Modellen?	51
3.4.2	Merkmalsextraktion und Erschließung	52
3.4.3	Typen automatisch extrahierter Metadaten	57
3.5	Formalisierte Annotation im Linked Data	57
3.5.1	Web of Documents → Web of Data	58
3.5.2	Umsetzung	59
3.5.3	Ausgewählte Modelle und Vokabulare	61
3.5.4	Stand der Forschung / Verwandte Arbeiten: Projekte und ihre Modelle	63
4	Metadaten für 3D-Modelle	65
4.1	Vorbemerkung	65
4.2	Namensräume	66
4.3	Entitäten	69
4.3.1	Gruppen	69
4.3.2	Einbindung externer Informationsquellen	70
4.3.3	Darstellung	71
4.4	Gruppe 1 – Modelle und ihre Instanzen	72
4.4.1	MODELINFO	72
4.4.2	MODELFILE	81
4.5	Gruppe 2 – involvierte Personen und Körperschaften	84
4.5.1	CONTRIBUTOR	84
4.5.2	PERSON	85
4.5.3	CORPORATE	86

4.6	Gruppe 3 –Ereignisse, reale Objekte, Orte und Konzepte	87
4.6.1	EVENT	88
4.6.2	OBJECT	89
4.6.3	PLACE	92
4.6.4	SUBJECT (CONCEPT)	94
4.7	Visualisierung	96
5	Bewertung und Ausblick	101
	Literaturverzeichnis	103
	Anhang	112
	Danksagung	134

Abbildungsverzeichnis

Abbildung 1:	3D-Gebäude-Entwurfsmodell und 2D-Planausgabe (Quelle: Institut AIDA, LU Hannover)	19
Abbildung 2:	Amorphe Freiform in Maya mit kompakter MEL-Geometriebeschreibung	25
Abbildung 3:	Per Laserscan aufgenommenes Objekt (Quelle: CGI, Uni Bonn)	26
Abbildung 4:	Verwendung von 3D-verarbeitenden Programmen (Auszug aus einer Befragung von Architekten, vgl. Abschnitt 2.4.3)	27
Abbildung 5:	3D-Gebäudemodell mit Visualisierung der im Gebäude enthaltenen strukturellen Komponenten (Quelle: Institut AIDA, LU Hannover)	29
Abbildung 6:	Verwendung von 3D-Modellen entlang des Prozesses der Planungsdetailierung	32
Abbildung 7:	Verwendungsszenarien extern erstellter 3D-Modelle	33
Abbildung 8:	Verwendung von 3D-Modellen in der Planungspraxis	34
Abbildung 9:	Gewünschte Suchmöglichkeit / 3D-Modelltyp	35
Abbildung 10:	Beziehung zwischen den Begriffen Wissen, Information, Daten und Zeichen nach (Rehäuser & Krcmar, 1996)	40
Abbildung 11:	Polygonales Modell im Wavefront .obj-Format, geöffnet im CAD- sowie im Texteditor	51
Abbildung 12:	Modellprozessierung (Teil 1) - Vorverarbeitungsschritte, Normalisierung, Objektabstraktion	54
Abbildung 13:	Modellprozessierung (Teil 2) - Segmentierung, Umwandlung sowie anschließende Merkmalsextraktion und Erschließung	55
Abbildung 14:	Schematische Darstellung der wichtigsten Entitäten und Relationen des 3D Metadatenmodells	69
Abbildung 15:	3D-Gebäudemodell eines real existierenden Bauwerks mit dargestellter Verortung (Quelle: PROBADO 3D)	93
Abbildung 16:	Browsing über 3D-Modelle mittels Begriffen, die durch automatisierte Klassifikation vergeben wurden. Anzeige der Konfidenz der Klassifikation am jeweiligen Modell (PROBADO 3D)	96
Abbildung 17:	Aussagen über ein 3D-Modell und dessen Kontext	98
Abbildung 18:	Mögliche Visualisierung der angereicherten Modellinformationen	99

Abbildung 19:	parabolisches Browsing über integrierte Klassifikation (PROBADO 3D)	125
Abbildung 20:	Tabellen der relationalen Datenbank	127
Abbildung 21:	Mittels MINT-Mappingtool realisierte Vorschau auf eine Detailansicht der Modellmetadaten in der Europeana	128

Abkürzungsverzeichnis

BIM	Building Information Modeling
CAD	Computer Aided Design
CAM	Computer Aided Manufacturing
CC	Creative Commons
CSG	Constructive Solid Geometry
DOI	Digital Object Identifier
HTML	Hypertext Markup Language
IFC	Industry Foundation Classes
NURBS	Non-uniform rational Bézier-Spline
OAI	Open Archives Initiative
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
PDF	Portable Document Format
RDF	Resource Description Framework
RDFS	Resource Description Framework Schema
SQL	Structured Query Language
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
W3C	World Wide Web Consortium
WWW	World Wide Web
XML	Extensible Markup Language

Einleitung und Zielsetzung

„Durch die Verknüpfung von Daten in verschiedenen Kontexten werden Zusammenhänge offenbar, welche zuvor nicht unmittelbar ersichtlich waren.“¹

Mit dem fortschreitenden Einfluss der Informationstechnologie hat sich die Art, wie Informationen repräsentiert werden, von einer analogen in eine digitale Darstellung verschoben. Darüber hinaus werden bereits vorhandene analoge Daten zunehmend digitalisiert. Architekturrelevante Darstellungen treten zunehmend als 3D-CAD-Modelle und 2D-Zeichnungsdateien, Building Information Models (BIM), Laserscan-Aufnahmen, digitale Bilder oder Videos auf. Da wissenschaftliche Bibliotheken und Archive den Auftrag haben, diese Materialien systematisch zu sammeln, zu erschließen, zu archivieren und optimal zur Verfügung zu stellen, sodass sie für neue planende und wissenschaftliche Arbeiten verwendet werden können, müssen sie konsequenterweise in der Lage sein, neue Medienformate zu akzeptieren.

Digitale Datenrepräsentation bedeutet nicht nur eine Änderung der Art und Weise, wie Inhalte gespeichert und verfügbar gemacht werden, sie revolutioniert auch die Verfahren, mit denen die Daten durchsucht und abgerufen werden können. Im analogen Zeitalter waren die Möglichkeiten, nach einem Buch mit spezifischen Inhalten zu suchen, beschränkt; es gab nur den Bibliothekskatalog mit seinen „Records“ (Einträge im Zettelkatalog). Auf der Suche beispielsweise nach einem bestimmten Buch im Bibliotheksbestand half es nicht, nur eine bestimmte Zeile zu kennen. Im fortschreitenden digitalen Zeitalter wäre eine solche Suche allerdings denkbar, denn es wird eine immer größer werdende Menge an Büchern gescannt. Zunächst führt das zu einer großen Menge digitaler Bilder, die die einzelnen Seiten darstellen. Im nächsten Digitalisierungsschritt hebt eine Optical Character

¹ Frei nach Christian Horchert, Open Data Network. Im Original heißt es: „Wenn man Daten in verschiedenen Kontexten miteinander verknüpft, erkennt man Zusammenhänge, die man vorher nicht hatte.“, vgl. <http://www.heise.de/newsticker/meldung/Big-Data-Rohstoff-der-Informationsgesellschaft-1659622.html>, abgerufen Dezember 4, 2012.

Recognition Software (OCR-Software) die „Suppen“ aus Pixeln durch die Identifizierung von Zeichen und Wörtern auf eine neue semantische Ebene. Erstellt man nun einen Volltextindex aus allen verfügbaren Büchern der Bibliothek, kann das Buch, aus dem man nur eine einzige Zeile kennt, gefunden werden. Der mit dem Semantic-Web-Paradigma² einhergehende Wandel ermöglicht zudem die Erkennung und Vernetzung semantischer Entitäten³ und damit die Einbeziehung zusätzlicher Informationsquellen außerhalb des „Silos“ Bibliothekskatalog, sodass man zusätzliche, nicht im Bibliothekskatalog vorhandene Informationen zum gesuchten Buch erhalten kann.

Trotz des Paradigmenwechsels von analogen zu digitalen Darstellungen in architektonischen Entwürfen und Konstruktionen basieren die Zugriffsmechanismen der aktuellen Bibliothekssysteme, wie sie derzeit für neue Medienformate wie 3D-Modelle praktiziert werden, noch auf einem einfachen, sich aus den wenigen seitens der Modellbereitsteller mitgelieferten Annotationen speisenden Metadaten-„Records“, was vergleichbar ist mit der auf einem Zettelkatalog basierenden Suche im analogen Zeitalter. Aus Sicht des Benutzers führt dies zu einem Informationsverlust, da weder die semantische Ebene innerhalb der Modelle noch der Modellkontext für eine textbasierte Suche berücksichtigt wird.

Methoden der Informationsextraktion sind bei 3D-Modellen jedoch im Vergleich zu Textdokumenten (mit Buchstaben als wohldefinierte Informationseinheiten) diffiziler, denn während durch OCR der Text explizit dargestellt wird, sind die informationstragenden Einheiten bei 3D nicht hinreichend diskret vorhanden. So müssen zunächst syntaktisch interpretierbare und erst in einem zweiten Schritt semantisch unterscheidbare Einheiten (wie beispielsweise Stockwerke von mehrgeschossigen Gebäuden) extrahiert werden. Für die Vernetzung dieser Einheiten mit zusätzlichen Informationsquellen ist ferner die Identifikation weiterer domänenspezifischer Einheiten mit inhaltlicher Relevanz notwendig (beispielsweise

² vgl. Ausführungen zur Genese des „Linked Data Web“ in den Abschnitten 3.3.1 sowie 3.5.1

³ Objekte mit eindeutigen Identitäten, beispielsweise Personen, Orte, Ereignisse, Konzepte. Identifizierbar mit URIs, vgl. 3.5.

Akteure, Konzepte, Objekte, Ereignisse), um eine explorative Suche über den Modellkontext und darüber hinaus zu ermöglichen. Der Ansatz der Vernetzung von Objekten über ihren Kontext beinhaltet zudem heuristisches Potential (vgl. das einleitende Zitat). Obwohl also Information vorhanden und extrahierbar ist sowie verknüpfbar gemacht werden kann, findet die Verwertung dieses Potenzials im Hinblick auf 3D-Architekturmodell derzeit in wissenschaftliche Bibliotheken und Archiven nur ansatzweise statt.

Diese Arbeit geht der Fragestellung nach, wie Information, welche aus 3D-Architekturmodellen gewonnen wird, in einem angesichts sich dynamisch entwickelnder Extraktionsmethoden flexiblen und erweiterbaren Modell abgebildet werden kann, und mit welchen Ansätzen die im Modell enthaltene Information mit weiterer Information vernetzt werden kann.

Hierzu werden die sich aus 3D-Modellierung, technischen Implikationen sowie inhaltlichen Vorgaben der Planung und Suche ergebenden Anforderungen erörtert. Im Weiteren werden die unterschiedlichen Ebenen der Informationsextraktion untersucht, um daraus Metadatatypen abzuleiten. Außerdem werden aktuelle Möglichkeiten der Wissensrepräsentation erörtert, insbesondere bezüglich Möglichkeiten der kontextualisierten Repräsentation von Inhalten.

Die Arbeit gliedert sich in die folgenden Abschnitte:

- Im ersten Kapitel werden die aktuellen Herausforderungen digitaler Bibliotheken aufgelistet und Handlungsfelder identifiziert.
- Das zweite Kapitel widmet sich digitalen 3D-Architekturmodellen, ihren technisch-strukturellen Attributen und ihrer Anwendung.
- Das dritte Kapitel widmet sich der Informationsextraktion, insbesondere aus 3D-Modellen, und der Wissensrepräsentation.
- Das vierte Kapitel beschreibt das im Hinblick auf die innerhalb der Kapitel 2 und 3 identifizierten Anforderungen entwickelte Metadatenmodell sowie eine exemplarische Visualisierung zu einem Beispielmmodell und dessen Kontext, wobei avancierte Fragen der graphen- / RDF-basierten Modellierung nicht behandelt werden.

- Eine abschließende Diskussion findet in Kapitel 5 statt.

1 Digitale Bibliotheken

1.1 Definition

Was intellektuell oder künstlerisch von Menschen erdacht und geschaffen wird, manifestiert sich in der Regel in Medien, z.B. als Buch oder Bild, und wird auf diese Weise an die Umgebung übermittelt. Informationseinrichtungen wie Bibliotheken, Archive oder Museen⁴ haben seit jeher die Aufgabe, solche Informationsobjekte⁵ strukturiert zu sammeln und mit ihrer Erfassung die Wiederauffindbarkeit zu gewährleisten.

Digitale Bibliotheken sind organisierte, dauerhafte Sammlungen von z.T. global verteilten digitalen Informationsobjekten, die ergänzt werden können um Technologien und Methoden zum Suchen von Informationen und zur Speicherung, Sicherung, Darstellung, Erstellung und Nutzung der Informationsobjekte, vgl. hierzu die Übersicht zu Aufbau und Funktionen digitaler Bibliotheken in (Endres & Fellner, 2000). Das DELOS Digital Library Reference Model (Candela u.a., 2011, S. 171) definiert digitale Bibliotheken wie folgt:

“An organisation, which might be virtual, that comprehensively collects, manages and preserves for the long term rich *Information Objects*, and offers to its *Actors* specialised *Functions* on those *Information Objects*, of measurable quality, expressed by *Quality Parameters*, and according to codified *Policies*.“

Digitale Inhalte können genuin digital oder durch Transformation analoger Objekte, z.B. Scanning, erstellt worden sein. Die meisten digitalen Bibliotheken sind bislang weitgehend auf klassische Publikationen und Informationsbestände beschränkt, die in textueller Form vorliegen, beispielsweise Zeitschriftenartikel oder Konferenzbände.

⁴ werden auch unter dem Begriff „GLAM“ institutions (galleries, libraries, archives and museums) subsumiert

⁵ Die Begriffe „Ressource“ und „Informationsobjekt“ werden in dieser Arbeit synonym verwendet.

1.2 Herausforderungen

Die Informationslandschaft befindet sich seit der Entwicklung des WWW zu Beginn der 90er Jahre und seinem darauffolgenden starken Wachstum in einer Umbruchphase. Die Möglichkeiten des digitalen Publizierens und Reproduzierens verweisen auf einen Paradigmenwechsel, der häufig mit der Ablösung der Handschrift durch den Buchdruck verglichen wird, siehe beispielsweise (Gradmann, 2012a). Auch David Stuart (Stuart, 2011) betont diesen Prozess, der sich in einer Verschiebung weg von gedruckten Informationen und hin zu „Born-digital“- oder „Nur-Daten“-Formaten manifestiert.

Die Herausforderungen und die sich aus dem Paradigmenwechsel ableitenden Handlungsfelder für wissenschaftliche Informationseinrichtungen, insbesondere für digitale Bibliotheken, werden im Folgenden skizziert.

1.2.1 Informationsvolumen

Die oben angesprochene Expansion des WWW ging zunächst einher mit der Zunahme an digital verfügbaren Informationen; laut einem IDC-Whitepaper war das Volumen im Jahr 2011 zehnmal größer als im Jahr 2006 (Gantz, 2008). Das reine Größenwachstum in Bytes ist insbesondere den nichttextuellen Dokumenten geschuldet⁶, vornehmlich den audiovisuellen Medien (AV-Medien)⁷.

John Naisbitt spricht in seinem Buch *Megatrends* (Naisbitt, 1982) folgendes Problem an:

“We are drowning in information but starved for knowledge.”

⁶ vgl. „Weltweiter IP-Datenverkehr steigt von 2010 bis 2015 um das Vierfache“, in http://www.cisco.com/web/AT/assets/docs/presse/11_Cisco_PA_VNI_FIN_0614.pdf, Prognose von Cisco Systems vom 14.06.2011

⁷ vgl. http://www.wired.com/images_blogs/business/2011/05/SandvineGlobalInternetSpringReport2011.pdf. Hinsichtlich der Zunahme der AV-Medien ist allerdings anzumerken, dass Medien mit wissenschaftlicher Relevanz nur einen geringen Teil der genannten Menge ausmachen. Sämtliche verfügbaren Daten zu sammeln, zu erfassen und zugänglich zu machen ist mit traditionellen bibliothekarischen Methoden, vgl. 3.2.2, nicht möglich und auch nicht Aufgabe von wissenschaftlichen Informationseinrichtungen.

Naisbitt verweist hier eingängig auf das sich aus dem gestiegenen Informationsvolumen ableitende Handlungsfeld für wissenschaftliche Informationseinrichtungen, mit Hilfe geeigneter Strukturierung und Indexierung dem Benutzer Werkzeuge in die Hand zu geben, damit er aus der verfügbaren Flut an Informationen relevantes Wissen destillieren kann.

Ein aktuell in Bibliotheken, so etwa an der Technischen Informationsbibliothek Hannover (TIB), praktizierter Versuch, die zunehmende Menge an Daten zu erschließen, ist der Rückgriff auf aus der Forschung zur Künstlichen Intelligenz (KI) bekannten Methoden, z.B. maschinelles Lernen zur automatischen Klassifikation, vgl. Abschnitt 3.4.2. Neben einer automatisierten Datenerfassung ist die Erschließung von Inhalten durch vernetzte Nutzergruppen ein in der informationswissenschaftlichen Literatur der letzten Jahre vieldiskutiertes Thema, vgl. (van Harmelen, 2006, S. 4) oder als Überblick (Peters, 2009).

Ein althergebrachtes bibliothekarisches Selbstverständnis ist das der qualitativen Auslese. Dabei erhalten die im Bibliotheksportal durchsuchbaren Inhalte entsprechend dem Erwerbungsprofil der jeweiligen Bibliothek und dem subjektiven Ermessen des zuständigen Fachreferenten das Prädikat „wissenschaftlich relevant“ und werden katalogisiert bzw. sacherschlossen. Mit der Informationszunahme ist ein solcher Weg ressourcentechnisch immer weniger gangbar, selbst unter Berücksichtigung der teilautomatischen Erschließung. Die Inhalte innerhalb dieser „Silos“, nämlich der einzelnen Bibliothekskataloge, werden traditionell nach dem analog geprägten Katalogparadigma, also dem Nachweis im Zettelkatalog, organisiert. Im Zeitalter vernetzter Informationen wird dem Benutzer somit ein reduzierter Suchraum mit „wissenschaftlich relevanten“ Inhalten zur Verfügung gestellt.

1.2.2 Erweiterung des Suchraums

Wachsende Anforderungen an digitale Bibliotheken werden durch die Benutzer vorgegeben; sie beruhen auf ihren allgemeinen Interneterfahrungen hinsichtlich eines schnellen, ortsunabhängigen und barrierefreien Informationszugriffs durch die Benutzung interaktiver Werkzeuge und ausgefeilter Suchmaschinentechnologien bzw. auf ihrem Umgang mit kollaborativen Diensten und Netzwerken.

Überdies stellen immer mehr wissenschaftliche Zeitschriften bibliografische Daten wie Inhaltsverzeichnisse, Abstracts und Keywords frei ins Internet, vgl. (Chen, 2010b). Konkurrenten, insbesondere Google mit Produkten wie „Google Scholar“ oder „Google Books“, die klar auf Kernaufgaben wissenschaftlicher Informationsdienstleister zielen, erlangen zunehmend Relevanz in der Informationsversorgung, vgl. eine mit „Google Scholar’s Dramatic Coverage Improvement Five Years after Debut“ überschriebene empirische Studie von (Chen, 2010a). Diese Angebote verfügen mit der Volltextsuche und der Möglichkeit zur explorativen Suche – also einer Suche, welche die durch Links hergestellten Beziehungen zwischen den Quellen zur Navigation benutzt – über einen entscheidenden Vorteil gegenüber der Recherche in traditionellen OPACs oder in anderen bibliothekarischen Online-Portalen.

Der Suchraum von Internet-Suchmaschinen wie Google umfasst sämtliche im Internet frei verfügbaren Informationsquellen, er ist also nicht wie in Bibliotheken auf einen Zettelkatalog und damit auf die bibliothekarisch erfassten Medien beschränkt. In diesem Zusammenhang weist ein OCLC-Bericht darauf hin, dass nur 1 % der in einer Studie Befragten bei Recherchen nach elektronischen Informationen auf Webseiten oder Portalen einer Bibliothek starten (De Rosa u.a., 2005, S. 1–17). Die W3C Library Linked Data Incubator Group⁸ merkt an (W3C Incubator Group, 2011):

“Library data today resides in databases which, while they may have Web-facing search interfaces, are not deeply integrated with other data sources on the Web. There is a considerable amount of bibliographic data and other kinds of resources on the Web that share data points such as dates, geographic information, persons, and organizations. In a future [...] environment, all these dots could be connected.”

⁸ <http://www.w3.org/2005/Incubator/lld/>

Um einem Bedeutungsverlust entgegenzuwirken, müssten Bibliotheken, so die Forderung von (Coyle, 2010), ihre Daten in die offene Architektur des World Wide Web integrieren:

“Today, we face another significant time of change that is being prompted by today’s library user. This user no longer visits the physical library as his primary source of information, but seeks and creates information while connected to the global computer network. The change that libraries will need to make in response must include the transformation of the library’s public catalogue from a stand-alone database of bibliographic records to a highly hyperlinked data set that can interact with information resources on the World Wide Web. The library data can then be integrated into the virtual working spaces of the users served by the library.”

Vor diesem Hintergrund resümiert (Stuart, 2011, Kapitel 7):

“Ignoring the web of data would seem to be the beginning of the end of librarianship.“

1.2.3 Medienformate

Seit einigen Jahren steigen Heterogenität und Komplexität von Medienformaten, die im wissenschaftlichen Kontext produziert werden. Daten werden von Instrumenten erfasst, mittels Applikationen erstellt oder durch Simulatoren erzeugt, es entsteht durch Software verarbeitete Information, die durch Wissenschaftler analysiert, transformiert oder weiterverarbeitet wird. (BLK, 2006, S. 8) kommentiert diesen Prozess wie folgt:

„Je stärker wissenschaftliches Arbeiten auf Informations- und Kommunikationstechnologien ausgerichtet ist, desto eher bringen wissenschaftliche Erkenntnisse neue Ausdrucksformen jenseits von Text und Bild hervor.“

Diese Diversifizierung von Medienformaten schlägt sich in unterschiedlichen Charakteristiken⁹ und Lebensdauern nieder, was insbesondere Auswirkungen auf die Erschließung, Bereitstellung und (Langzeit-)Archivierung hat.¹⁰ Jim Gray benennt diese datenintensive Forschung als ein neues viertes Paradigma einer komplexer werdenden Wissenschaft (Hey u.a., 2009). Der Abschlussbericht der High Level Expert Group on Scientific Data (Wood, John, 2010) stellt vor diesem Hintergrund fest:

“Knowledge is power; Europe must manage the digital assets its researchers generate.“

und fordert damit Antworten auf drängende Fragen zum Management wissenschaftlich generierter und konsumierter Daten und Informationen.

1.2.4 Strukturelle Aspekte / Dokumentenmodelle

Digitale Dokumente imitieren aktuell noch analoge Techniken, z.B. elektronische Zeitschriften im PDF- oder HTML-Format in Anlehnung an die Printausgabe, und können als digitale Abbildungen von Bibliotheksbeständen mit erweiterten Navigationsmöglichkeiten verstanden werden. Sie werden aber im Zuge des eingangs unter 1.2 angesprochenen Paradigmenwechsels zunehmend durch dynamische Publikationen sowie miteinander verlinkte textuelle wie nichttextuelle Objekte erweitert oder gar abgelöst, vgl. (Nowakowski u.a., 2011), was auch einen begrifflichen Wandel bedeutet. So hat das WWW zu einer veränderten Sicht auf Dokumente geführt, und man habe es laut (Gradmann, 2012a) in den WWW-genuinen

⁹ Die unterschiedlichen Repräsentationsformen z.B. von 3D-Modellen steigen weiterhin an, trotz der Bemühungen, allgemeingültige Standards und Interoperabilität zu definieren, siehe BIM / IFC, Abschnitt 2.3.3.

¹⁰ Mit CAD erstellte digitale 3D-Modelle sind besonders problematisch für Bibliotheken und Archive, da sie sehr volatil sind, sich auf eigene mathematische Algorithmen stützen, um Formen und Strukturen zu repräsentieren und verpackt sind in komplexe, meist proprietäre und sich schnell entwickelnden Software-Produkte, die teuer, digital verschlüsselt und innerhalb von Jahren veraltet sind.

Systemen statt mit „Dokumenten“ mit „Aggregationen“ zu tun; und an die Stelle der „Suche“ trete die „Exploration“.¹¹

Die teilweise automatische Erkennung von objektinhärenter Semantik¹² – allgemein als *inhaltsbasierte Erschließung* bezeichnet – vergrößert die Menge an verfügbaren informationstragenden Entitäten. Dazu gehört auch die z.T. hochkomplexe Binnenstruktur von 3D-Modellen, vgl. 2.3. (Gradmann, 2009) fasst dies wie folgt zusammen:

„Die Anzahl der diskreten, digital prozessierbaren und semantisch unterscheidbaren Einheiten steigt also durch [...] und zugleich drastisch gesteigerte Granularität gleich um mehrere Größenordnungen.“

Beide Typen von Informationsobjekten, sowohl die aggregierten Objekte als auch die inhärenten Entitäten, erfordern Instrumente der Organisation und eine aufbereitete Visualisierung, um für den Benutzer einen Mehrwert darzustellen.

1.3 Fazit

Die oben genannten Herausforderungen zwingen digitale Bibliotheken zukünftig zu einer Erweiterung ihres Bestandes um neue Dokumentformen, z.B. 3D-Modelle, zu einer Anpassung der bibliothekarischen Prozesskette für die Handhabung dieser Dokumentformen und zu einer strukturierten, kontextualisierten Organisation ihrer gesamten Informationsressourcen. (Gradmann, 2009) formuliert dies wie folgt:

„Angesichts dessen benötigen wir neue Werkzeuge, um große Mengen digitaler Informationsobjekte einschließlich ihrer Konstituenten gedanklich zu organisieren und können dabei nicht mehr auf das althergebrachte bibliothekarische Katalogparadigma zurückgreifen. Wir müssen Mittel finden,

¹¹ vgl. Diskurs zum Dokumentenbegriff im digitalen Zeitalter, z.B. unter (Voß, Jakob, 2009) oder (Frohmann, 2009)

¹² „Semantik“, allgemein als Wissenschaft von der Bedeutung der Zeichen oder Symbole definiert, wird analog auch im Kontext dieser Arbeit angewendet, hier also: bedeutungstragende, unterscheidbare Einheiten innerhalb von Modellen.

bislang unvorstellbare Aggregationen strukturierter digitaler Informationen zu verstehen oder zumindest verstehbar zu machen.“

Eine wesentliche Rolle in der Öffnung und Vernetzung von Daten spielt das Konzept Linked Open Data (LOD)¹³, vgl. Einführung in Abschnitt 3.5, das eine domänenübergreifende Vernetzung von frei im Internet verfügbaren Daten ermöglicht. Die zentrale Idee von LOD besteht darin, dass Informationsressourcen aus dem Wissenschaftsbereich und darüber hinaus aufeinander referenzieren, in ihrem jeweiligen Kontext visualisiert werden können und, u.a. mittels semantischer Suchtechnologien, wie in einer einzigen verteilten Datenbank abgefragt werden können, vgl. (Heath & Bizer, 2011). Tim Berners-Lee statuiert in (Berners-Lee, 2006b), dass in Kontextualisierung sowie neuen, vielfach unerwarteten Verwendungsmöglichkeiten der Daten das Potenzial liege, für die Benutzer Mehrwerte zu schaffen.

Sowohl Microsoft mit „Microsoft’s Satori“ als auch Google mit „Google Knowledge Graph“ setzen auf Linked Data. Dabei werden aus Webseiten unstrukturierte Informationen extrahiert, um daraus automatisch Entitäten wie Orte, Begriffe oder Personen zusammen mit ihren Relationen strukturiert abzubilden. Auf diese Weise wird eine quasi ontologiebasierte Suche ermöglicht, die das Potenzial einer Volltextsuche erheblich übertrifft, vgl. (Gallagher, 2012).

Das Feld der Strukturierung und Organisation von Wissen ist seit jeher die Hauptqualifikation von Bibliotheken, vgl. 3.2. So ist auch seitens (insbesonderer wissenschaftlicher) Bibliotheken seit einigen Jahren ein zunehmendes Interesse am Thema Linked Data zu verzeichnen (Isaac, 2010), und es werden immer mehr

¹³ Linked Data erlaubt technisch sowohl geschlossene Anwendungen als auch das offene Netz der Linked Open Data, das auch als Linked Open Data Cloud (LOD) bezeichnet wird. Im Folgenden wird einheitlich von „Linked Data“ gesprochen, da insbesondere das Thema Vernetzung und Kontextualisierung von Information im Fokus der Arbeit steht.

bibliothekarische Daten als LOD zur Verfügung gestellt¹⁴, deren Bedeutung in (W3C Incubator Group, 2011) hervorgehoben wird.

(Isaac, 2010, S. 7) hebt im Kontext der zweiten SWIB¹⁵-Konferenz die Kompetenz und Erfahrung von Bibliotheken in der „Produktion hochqualitativer Metadaten“ als wichtige Elemente des Linked Data hervor.

¹⁴ Wichtige Akteure aus dem Bibliotheksbereich, die Daten als Linked Open Data zur Verfügung stellen, sind die Library of Congress mit den LoC Subject Headings (LCSH) und die Deutsche Nationalbibliothek mit der Gemeinsamen Normdatei (GND).

¹⁵ „Semantic Web in Bibliotheken“, jährliche wechselnd in Hamburg (ZBW) und Köln (hbz) stattfindende Fachtagung, erstmalig in 2009 und zunächst für die deutschsprachige Community, vgl. <http://www.swib09.de/>; seit 2012 englischsprachig und international ausgerichtet.

2 3D-Modelle

2.1 Definition

Modelle sind ein elementares Werkzeug zum Umgang mit der realen Welt und zum Verstehen von komplexen (z.B. räumlichen) Zusammenhängen. Jedes Schema, jede Abbildung ist ein Modell von etwas, und so bezeichnen (Reichle, Siegel, & Spelten, 2008) Modelle als einen „Bildkörper unseres Wissens“, dessen „Interpretation [...] immer auch eine Interpretation unserer eigenen Wahrnehmung“ ist. Das Konzept der Modellbildung ist so geläufig, dass Modelle, da sie auch flüchtigen, komplexen oder abstrakten Dingen eine Gestalt geben können, sich mitunter über die Realität stellen bzw. als die Realität ersetzend wirken können; dies gilt insbesondere für metaphorische Modelle mit besonderer visueller Kraft und Suggestivität, vgl. (Gradmann, 2012b).

Modelle werden gemeinhin als Abbilder von oder als Vorbilder / Entwürfe für Systeme oder Objekte verstanden. Stimmen Modelle optisch oder strukturell mit dem Referenzobjekt überein, werden sie gemeinhin leichter als Modelle wahrgenommen.

2.1.1 Original und Modell

Ein Modell entspricht nie gänzlich dem Original, da die Modellbildung mit einer Transformation verbunden ist, sodass Attribute des Originals durch Abstraktion wegfallen oder das Modell zusätzliche Attribute aufweist, die nicht durch das Original bestimmt sind. Dieser Mechanismus wird von (Stachowiak, 1973) anhand dreier Merkmale charakterisiert, nämlich Abbildung, Verkürzung und Pragmatismus:

„Modelle sind stets Modelle von etwas, nämlich Abbildungen, Repräsentationen natürlicher oder künstlicher Originale, die selbst wieder Modelle sein können. [...] Modelle erfassen im Allgemeinen nicht alle Attribute des durch sie repräsentierten Originals, sondern nur solche, die den jeweiligen Modellerschaffern und / oder Modellbenutzern relevant scheinen. [...] Modelle sind ihren Originalen nicht per se eindeutig zugeordnet. Sie erfüllen

ihre Ersetzungsfunktion 1) für bestimmte – erkennende und / oder handelnde, modellbenutzende – Subjekte, 2) innerhalb bestimmter Zeitintervalle und 3) unter Einschränkung auf bestimmte gedankliche oder tatsächliche Operationen.“

Zusammengefasst lässt sich festhalten: Modelle repräsentieren stets ein Original, das wirklich vorhanden, geplant oder fiktiv ist. Die Realität wird entsprechend der subjektiven Wahrnehmung des Modellkonstruktors abstrahiert. Ein Modell erfasst nicht jegliche Attribute des Originals, sondern einen Ausschnitt. Modelle können unter bestimmten Bedingungen und bezüglich bestimmter Fragestellungen das Original ersetzen.

Modelle lassen sich unterscheiden in Bild- oder Darstellungsmodelle als anschauliche Abbildungen des Originals, in semantische Modelle, die Wahrnehmungen, Vorstellungen, Gedanken oder Begriffe umsetzen, und in tendenziell technische Modelle als dreidimensionale Repräsentationen – z.B. 3D-Modelle aus der Physik, der Medizin oder der Architektur. Auf letztere soll im weiteren Verlauf näher eingegangen werden.

2.1.2 Anwendungskontexte → Architektur

(Stachowiak, 1973) unterscheidet bei wissenschaftlichen Modellen zwischen didaktischen Modellen zur Veranschaulichung von Zusammenhängen, Experimentalmodellen zur Ermittlung und Überprüfung von Hypothesen, theoretischen Modellen zur Vermittlung von Erkenntnissen über Sachverhalte und operativen Modellen zur Bereitstellung von Entscheidungs- und Planungshilfen. Zur letzten Kategorie gehören insbesondere 3D-Modelle der Architektur. Architekten sind in ihrer Wahrnehmung stark visuell orientiert, daher greifen sie häufig auf Modelle zurück, z.B. zur Verdeutlichung des Tragverhaltens einer Konstruktion, dessen statische Wirkungsweise ohne Modell gedanklich erheblich schwerer zu erfassen wäre.

Eines der inhaltlich wichtigsten Unterscheidungsmerkmale von Modellen ist, ob sie als deskriptive, abbildungsorientierte Modelle dazu eingesetzt werden, ein existierendes reales Objekt zu beschreiben, oder ob sie als präskriptive, konstruk-

tionsorientierte Modelle dazu dienen, auf ihrer Grundlage ein Objekt herzustellen. Die Unterscheidung wirkt sich auf die Kontextualisierung aus, da für deskriptive und präskriptive Modelle jeweils andere Kontexte relevant sind, z.B. real existierende Gebäude bei deskriptiven Modellen und Wettbewerbe bei präskriptiven Modellen.

3D-Modelle der Architektur können deskriptiven Charakter besitzen, beispielsweise als dreidimensionale Repräsentationen antiker Rekonstruktionen oder bestehender Bauwerke, als Teil einer Ausschreibung in Form von Objektmodellen bestimmter Bauteile oder als Objektrepräsentationen in aktuellen Produktkatalogen in Form von 3D-Möbelmodellen.

Allerdings stellen in der Architektur als einer planenden Disziplin die präskriptiven 3D-Modelle naturgemäß den Hauptanteil. Beispiele sind detaillierte Konstruktionsmodelle eines Brückenentwurfs oder Modelle, die bei Wettbewerben für die Entwurfsvisualisierung herangezogen werden.

(Boeykens & Bogani, 2008) führt analog dazu unterschiedliche „levels of reality“ für die Spezifikation der Beziehungen zwischen Objekten der realen Welt und 3D-Modellen auf, soweit sie existieren¹⁶:

“- *Real object*, which exists as a tangible subject;

- An *historical* object, such as a reconstruction of a demolished or perished building. It could also pertain to the recreation of a real building in a particular phase in time, e.g. before a certain restoration phase.”

bzw.

¹⁶ “However, there is the possibility that this 3D Model lacks a counterpart in the real world. An imaginary building could belong to a design studio assignment or could be created from the imagination of a graphics artist. Another example is the model of a particular chair, from the catalog of a manufacturer. It represents a specific model, but there is no singular relationship with a particular real world chair. It might even be a generic product, to be used in a technical document or as part of a visualization, e.g. an unbranded chair.”

“- A *planned* or *designed* object, which could be realized. Typical examples include models for architectural competitions, from which usually only a single winning design will finally be realized as a real world object;

- An *imaginary* object, which could be almost anything, without the necessity of it being realized. This not only includes architectural design studio results, but even pure imagined models, such as a spaceship or a fantasy dungeon.”

Im Folgenden wird die Rolle von Modellen im architektonischen Planungsprozess untersucht.

2.2 Digitale Modellierung

Die Arbeit von Architekten umfasst heterogene und hauptsächlich multimediale Inhalte, der Arbeitsbereich reicht von technischen bis hin zu künstlerischen Anwendungen. Digitale Arbeits-, Produkt-, Kommunikations- oder Fertigungsmodelle haben innerhalb der letzten fünf Jahrzehnte sowohl die analogen, auf Papier basierenden Zeichnungen als auch die physischen Modelle aus ihrer planungs-, ausführungs- und dokumentationsunterstützenden Rolle verdrängt (ein Überblick über diese Entwicklung gibt 2.2.2). Zwar werden im Planungsprozess immer noch traditionelle Mittel wie Skizzen oder Modellbau angewandt, die endgültige Planung und Darstellung wird aber fast ausschließlich mit Hilfe von Computer Aided Design Programmen (CAD-Programme) erstellt. Computer Aided Design bedeutet wörtlich übersetzt „computerunterstütztes Entwerfen“ und bezeichnet die Unterstützung des Zeichen- und Planungsprozesses durch Computer. Wird mittels CAD-Programm eine dreidimensionale Zeichnung erstellt, spricht man von einem 3D-CAD-Modell. Im Folgenden wird der Kurzbegriff 3D-Modelle verwendet; alle weiteren Ausführungen sind auf diese digitalen Modelle bezogen.

Architekturmodelle lassen sich hinsichtlich dessen, was sie repräsentieren, einerseits unterteilen in Modelle von Gebäuden sowie von Städten und Umgebungen, andererseits in Modelle von Bauteilen und Objekten, z.B. Einrichtungsgegenständen. Bei Gebäude- sowie Stadt- und Umgebungsmodellen sind im Gegensatz zu den Objektmodellen insbesondere räumliche und strukturelle Aspekte maßgeblich.

Für Gebäudemodelle und Objektmodelle ist außerdem die jeweilige Form ein Unterscheidungs- und Identifikationskriterium. Die Übergänge von Gebäude zu Objekt können fließend sein, beispielsweise bei Installationsobjekten wie Pavillons. Eine weitere Mischform bilden darüber hinaus Szenenmodelle, die nur einen Zimmerausschnitt mit zwei bis drei Wänden modellieren; sie dienen insbesondere zur Perspektivenerzeugung.

2.2.1 Planungsphasen und Modelltypen

3D-Modelle werden im Zusammenhang mit der Planung und Herstellung von Bauwerken oder Objekten eingesetzt. Im Bauwesen existiert für die gedankliche Entwicklung – vornehmlich eines Bauwerks – der Begriff der „Planung“, in der Architektur zusätzlich der domänenspezifisch eher künstlerisch geprägte und bis zur Ausführungsplanung verwendete Begriff des „Entwerfens“¹⁷. Der Planungsprozess wird pragmatisch in Phasen gegliedert. So sieht analog die Vergütung eines Architekten oder Ingenieurs – in Deutschland in der Honorarordnung für Architekten und Ingenieure (HOAI) geregelt – insgesamt neun Leistungsphasen vor, vgl. (Hartmann, 2007). Diese Sequenz dient dem Ziel, Produktinformationen austauschbar sowie Planungsleistungen abrechenbar zu machen. Aus pragmatischen Gründen müssen auch die planungsunterstützenden Systeme des Computer Aided Design (CAD) entsprechend den HOAI-Phasen standardisierte Repräsentationen erstellen können, vgl. (Steinmann, 2004).

Das Entwurfsmodell ist das Arbeitswerkzeug des Architekten, es dient dazu, unterschiedliche Kubaturen zu testen und räumliche bzw. strukturelle Zusammenhänge im Planungs- und Entwurfsprozess zu antizipieren. Sie werden insbesondere zum Experimentieren eingesetzt, wenn analoge Methoden schnell an ihre Grenzen stoßen, z.B. bei der Ausgestaltung freier Formen, vgl. NURBS in Abschnitt 2.3.2. Das Modell ist zwar ein Instrument, um räumliche Zusammenhänge zu veranschaulichen, aber es abstrahiert eher, als detailliert die (künftige) Realität abzu-

¹⁷ Für einen Überblick über die historische Entwicklung der architektonischen Entwurfspraxis und ihre kultur- und erkenntnistheoretischen Grundlagen wird auf (Hassenewert, 2006) verwiesen.

bilden. Im städtebaulichen Kontext kommen vornehmlich Massenmodelle zur Anwendung, in denen Bauwerke als einfache, massive Körper repräsentiert werden. Wichtig bei diesen Modellen ist weniger die hohe Detailtreue als vielmehr die Darstellung der zentralen Ideen und des Konzeptes, des Entwurfs. Damit unterscheiden sie sich auch von den vorgenannten deskriptiven 3D-Modellen, z.B. Gebäuderekonstruktionen.



Abbildung 1: 3D-Gebäude-Entwurfsmodell und 2D-Planausgabe (Quelle: Institut AIDA, LU Hannover)

Das Entwurfsmodell bildet eine Diskussionsgrundlage und die Vorstufe zum Präsentationsmodell, das sorgfältiger ausgearbeitet ist und als Teil der Planungsdarstellung zur Präsentation des Entwurfs dient, z.B. für Bauherren oder Wettbewerbsauslober. Zweck der Entwurfsmodelle ist die schnelle Erzeugung zweidimensionaler Darstellungen, vgl. Abbildung 1. Sie können außerdem zur Ausgabe perspektivischer Darstellungen benutzt werden, oder sie dienen als Eingabe für Systeme zum 3D-Druck, dem sogenannten „Rapid Prototyping“.¹⁸

Produktmodelle dienen zur Durchführung von Kosten- und Strukturanalysen, u.a. in der Ausschreibungsphase von Bauleistungen. (Hovestadt, 1998) bezeichnet die Modelle der Planung und Ausführung als „Produkt-, Kommunikations- und Ferti-

¹⁸ 3D-Drucker gehören heutzutage zur normalen Ausrüstung universitärer Architektur-Modellbauwerkstätten. Eine 1:1-Fertigung, z.B. mit CNC-Maschinen, wird v.a. im Standard-Fertighausbau, zunehmend aber auch in Bereichen des Mass-Customization eingesetzt.

gungsmodelle, die über die Montage „externalisiert“, d.h. in ein physisches Gebäude überführt [werden].“

Während der Nutzungsphase dienen 3D-Modelle ferner als Grundlage für die Aufgaben der Gebäudebewirtschaftung und -steuerung, die ein physisches Gebäude über seine Standzeit begleiten, dessen Funktionalität grundlegend erweitern können und die Grundlage bzw. das Referenzmodell für das Facilities Management bilden, z.B. bei Wartung, Umbau oder sicherheitstechnischen Einbauten.

Jenseits der Errichtung oder dem Betrieb von Gebäuden ist ein weiteres wichtiges Anwendungsgebiet von Modellen der Einsatz in Wissenschaft und Lehre. Digitale Rekonstruktionen veranschaulichen unterschiedliche Aspekte und Dimensionen der Bewahrung von Kulturerbe (antike Objekte, historische oder zerstörte Gebäude, usw.). Hier spielen durch den Zuwachs des Informationsvolumens die inhaltlich-qualitativen Aspekte eine explizit starke Rolle, vgl. (Pfarr, 2010, S. 75).

2.2.2 IT-Entwicklung und Aspekte der Computerunterstützung im Planungsprozess

Der Beginn des Paradigmenwechsels weg von analogen, auf Papier basierenden Zeichnungen und physischen Modellen und hin zur Computerunterstützung von Planungsprozessen reicht bis in die Mitte der 1960er Jahre. Basierend auf Ideen des Computergrafik-Pioniers Ivan Sutherland begründen sich die Anfänge des Computer-Aided Design (CAD) in dem prototypischen, interaktiven Programm „Sketchpad“ (Ivan Sutherland, 1963). Sketchpad bietet mittels einer visuellen Schnittstelle bereits die Möglichkeit, Beziehungen zwischen geometrischen Grundelementen eingegebener 2D- und 3D-Geometriedaten festzulegen und zu manipulieren. In den 1970er Jahren liegt der Fokus auf der Entwicklung von Systemen, die den Gebrauch von manuellen Werkzeugen zum 2D-Zeichnen nachahmen und ergonomischen Notwendigkeiten der Entwerfer nachkommen. Parallel werden Computerwerkzeuge für nichtgraphische Aspekte des Entwerfens, wie dem Gebrauch von Datenbank-Management-Systemen (DBMS) für die quantitativ bestimmbare Aspekte von Gebäuden zu benutzen. Es folgt eine Suche nach Verfahren, die Erzeugung, Bewertung und Darstellung von Entwurfslösungen erlauben soll. Erstmals sollen auch kognitive Notwendigkeiten der Entwerfer er-

füllt werden, indem darauf fokussiert wird, wie die Informationen und das Wissen wahrgenommen, erworben, gespeichert und verarbeitet werden, vgl. (Mitchell, 1978). Aus der künstlichen Intelligenz (KI) entlehnte Methoden untersuchen die Kodierung des architektonischen Wissens in Designwerkzeugen (sogenannte „Knowledge into Design-Tools“) und gipfeln in Versuchen, Werkzeuge für die komplette Entwurfsautomatisierung zu entwickeln, deren Versprechen – analog zu denen der „starken KI“ jener Zeit – nicht erfüllt werden können. Ab den 1980er Jahren verschiebt sich der Fokus wieder auf die zunehmende dreidimensionale Computerunterstützung von Planungsprozessen, gefolgt von den ersten CAD-Produktmodellen in den 1990er Jahren.

Analog zur Strichzeichnung in Handarbeit besteht das frühe CAD aus Grundelementen wie Punkten, Linien, Bögen und Kreisen und aus Werten, z.B. Koordinaten, Längen, Winkel. Die Darstellungen von Bauteilen bestehen ebenfalls aus vielen einzelnen Grundelementen. Durch Gruppierung können die aus einzelnen Grundelementen bestehenden Bauteile als „Block“ behandelt, und damit also als Zeichnungsobjekte gelöscht, kopiert, bewegt, verzogen oder gespiegelt werden. Als Vorformen einer objektorientierten Systematik arbeitet die folgende Generation von CAD-Programmen mit sogenannten parametrisierten Blöcken, die in programminternen Modellbibliotheken organisiert und bei neuen Planungsaufgaben abgerufen werden können. Die Detaillierungsdarstellung dieser Objekte reagiert auf den Zeichnungsmaßstab und die Parameter (z.B. der Wandaufbau) sind in Attribut-Wert-Paaren einstellbar. Die aktuellen Programme des Building Information Modeling (BIM) sind durchgängig parametrisch und beinhalten strukturierte Information zu allen enthaltenen Objekten. Die konsequent 3D- / BIM- gestützte Planung bietet eine Reihe von Vorteilen, darunter die immersive Visualisierung des Bauvorhabens und automatisierte Konsistenzerhaltung unterschiedlicher Ansichten und Schnitte, eine automatisierte, präzise Mengenermittlung und die einfache Anbindung verschiedener Analyse- und Simulationstools. Problematisch ist die Erhaltung der Semantik durch Konvertierung in andere CAD-Formate als die

als Austauschformat standardisierten Industry Foundation Classes (IFC).¹⁹ BIM-Software integriert zusätzlich zu den drei geometrischen Abmessungen noch weitere, insbesondere Facilities Management -relevante Parameter, vgl. (Holzer, 2011), z.B. Zeit, Energiebilanz, Lebensdauer, Nachhaltigkeit, Kosten, Nutzung, etc.

2.3 Strukturelle Eigenschaften und technische Attribute

Im aktuellen Diskurs über die Erschließung und Archivierung nichttextueller Dokumente werden digitale 3D-Modelle in Abgrenzung zu z.B. Bildern oder Videos auch als „komplexe Objekte“ bezeichnet, vgl. (Bienert, 2012). Die folgende Übersicht fasst die strukturbedingten Attribute von 3D-Modellen zusammen, die z.T. bei der metadatenbasierten Beschreibung zu berücksichtigen sind, und zwar nicht nur als zusätzliche technische Attributfelder, sondern auch im Hinblick auf die Grundkonzeption des Metadatenmodells. Auf die im Folgenden diskutierten Herausforderungen wird in Kapitel 4 wieder zurückgekommen.

2.3.1 Anzahl an Dateien

Im Vergleich zu Bilddokumenten bestehen 3D-Modelle normalerweise aus mehr als einer Datei. Mit der geometrie- und strukturbeschreibenden Modelldatei sind oft mehrere Texturdateien (also Bilddateien) verknüpft, die auf die Modelloberfläche projiziert werden und ihr so eine Materialität verleihen. Texturen werden zwar nicht zwingend zur geometrischen Modelldarstellung benötigt, sind aber ebenfalls Träger semantischer Information, vgl. (Varma & Zisserman, 2009); so lässt – insbesondere bei ähnlicher Kubatur – erst die Textur auf den Modelltyp schließen. Die Verteilung von modellinhärenten semantischen Informationen auf mehrere Modell- und Bilddateien kann z.B. bei notwendigen, möglichst informationsverlustfreien Konvertierungsschritten im Zuge der Langzeitarchivierung eine Herausforderung darstellen.

¹⁹ <http://www.buildingsmart.de/>

Darüber hinaus können mit der Modelldatei und ihren Texturdateien weitere textuelle oder nichttextuelle Informationen wie Perspektiven und Videos oder Datenblätter mit Zusatzinformationen geliefert werden, die dem Benutzer ebenfalls zugänglich gemacht werden sollten. Diese Zusatzinformationen sind im Gegensatz zu den Texturen aber nicht direkt mit der Modelldatei verknüpft.

2.3.2 Geometrische Repräsentation

3D-Modelle können von dem jeweils benutzten Aufnahme- oder Modellierungswerkzeug in verschiedenen geometrischen Repräsentationsformen ausgegeben werden. Die Repräsentationsformen beeinflussen insbesondere die Dateiformate, vgl. 2.3.3. Die geometrische Repräsentation ist außerdem wichtig für das jeweils applizierte Shape-Retrieval-System, vgl. Kapitel 1. Man unterscheidet Oberflächenmodelle und Volumenmodelle, wobei Volumenmodelle (wenn es sich nicht um volumengetriebene Darstellungsformen wie Voxel-Modelle handelt, die in der Medizin verwendet werden) als eine Erweiterung der Flächenmodelle gelten, soweit sie ein Volumen mit einer geschlossenen Oberfläche (bezeichnet als „watertight“) umfassen. Insbesondere im 3D-Druck sind geschlossene Oberflächen Voraussetzung. Ein eindeutig dargestelltes Volumen erleichtert ebenfalls die Verarbeitung im Shape Retrieval, vgl. (Tangelder & Veltkamp, 2008, S. 4). In der Praxis gibt es für Oberflächen- wie für Volumenmodelle viele unterschiedliche Repräsentationstechniken. Eine vertiefende technische Darstellung findet sich bei (Klein, 1997, Kapitel 1 „Körper“) und bei (Fellner, 1992). Im Folgenden wird eine Übersicht über die im Architektur- und Konstruktionsbereich wichtigsten Techniken gegeben.

Polygonale Netze

Die meisten von CAD-Programmen erstellten und im WWW verfügbaren Modelle sind polygonale Netze. Eine polygonale 3D-Modellrepräsentation kann beispielsweise als Dreiecksnetz vorliegen, jedoch sind auch beliebige andere Polygone möglich. Bei der Architekturmodellierung werden meist Vierecke verwendet, da normalerweise rechtwinklige Bauteile wie Wände oder Fenster konstruiert werden. Polygonale Netze verbinden einzelne Punkte im dreidimensionalen

Raum, vorstellbar als n aneinandergefügte Ebenen, also lokal planare Einheiten, selbst dann, wenn mit dem Modell eine Kugel dargestellt wird. In der CAD-Modellierungspraxis kann eine als polygonales Netz repräsentierte Oberfläche fallweise, z.B. nach der Formatkonvertierung, sich schneidende Polygone und ggf. Löcher enthalten, sodass ein Volumen nicht eindeutig dargestellt wird.

Non-uniform rational Bézier-Spline (NURBS) and Non-rational Geometries

Insbesondere bei der Modellierung sogenannter „Non-Standard-Architecture“, vgl. (Migayrou u. a., 2003), wird auf architektonische 3D-Modellierungswerkzeuge zurückgegriffen, die auf Polynomen basierende NURBS-Geometrien herstellen können. Die zur Darstellung einer freigeformten Geometrie benötigte Beschreibung, siehe Abbildung 2, ist kompakter als eine Annäherung an die Form mittels n Flächen, woraus eine relativ lange Zeichenkette resultieren würde. Die architektonischen Anwendungsgebiete der ursprünglich aus der Filmindustrie und dem Produktdesign stammenden Programme sind neben der Modellierung komplexer Geometrien wie Tragwerke oder Freiformen die Lichtplanung / Lichtsimulation, Prototyping, Design, Animation und die klassische Visualisierung. Ein weiterer Vorzug in der Anwendung eines NURBS-Modelers liegt in einer höheren Flexibilität, wenn dieser mit dem verwendeten CAD-System zusammenarbeitet.

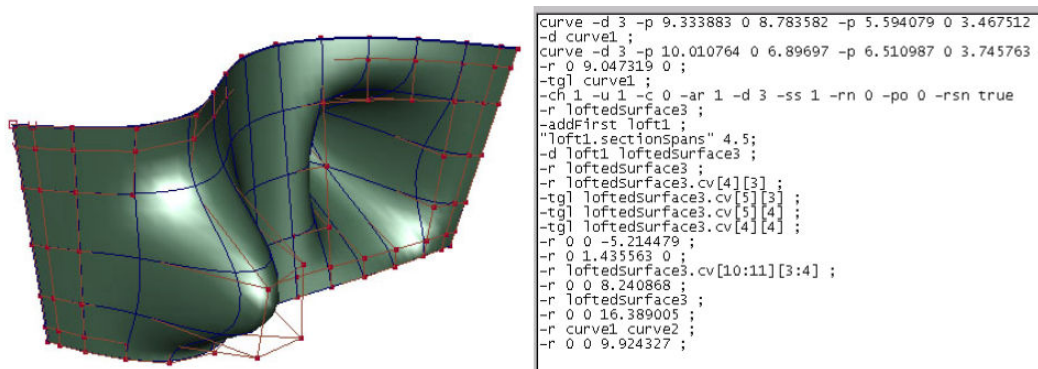


Abbildung 2: Amorphe Freiform in Maya mit kompakter MEL²⁰-Geometriebeschreibung

Constructive Solid Geometry (CSG)

Dabei wird das Modell aus der Vereinigung geometrischer Primitiva (Ebene, Kugel, Zylinder, Torus, Kegel etc.) aufgebaut; CSG wird insbesondere im Maschinenbau und im konstruktiven Ingenieurbau eingesetzt. Da sie per Konstruktion geschlossene Volumen erzeugen, sind CSG-Modelle immer Volumenmodelle.

Punktwolkenmodelle

Punktwolkenmodelle werden vornehmlich mittels 3D-Laserscanner aus real existierenden Objekten (z.B. als Gebäudeaufmaß oder Aufnahme von Kulturgütern) erzeugt. Sie bestehen im Gegensatz zu den vorgenannten linien- und flächenbeschreibenden Modellen aus einzelnen Punkten, die jeweils einen Punkt auf der gescannten Oberfläche des Objekts repräsentieren. Werden Teile des Objekts nicht oder nur unzureichend aufgenommen, führt dies zu Löchern im Punktwolkenmodell, die in einem weiteren Verarbeitungsschritt geschlossen werden müssen, z.B. durch Rekonstruktionsalgorithmen bzw. Interpolationsverfahren wie es z.B. (Kazhdan, Bolitho, & Hoppe, 2006) gezeigt haben. Eine weitere Herausforderung besteht darin, dass die gescannten Punkte in der Regel nie genau auf der Oberfläche lokalisiert sind und dass es Ausreißer in den Daten gibt, das sogenannte „Rauschen“, vgl. Abbildung 3. Punktwolkenmodelle, aber auch CSG- und NURBS-

²⁰ MEL steht für „Maya Embedded Language“. Hierbei handelt es sich um eine Skriptsprache innerhalb der 3D Grafiksoftware Maya der Firma Autodesk.

Modelle können in eine polygonale Darstellung überführt werden. Der Vorgang heißt „Tessellierung“, ist im Allgemeinen nicht eindeutig umkehrbar und hat eine einfache gemeinsame Modellrepräsentation zum Ergebnis, beispielsweise als Eingabe für das Shape Retrieval.

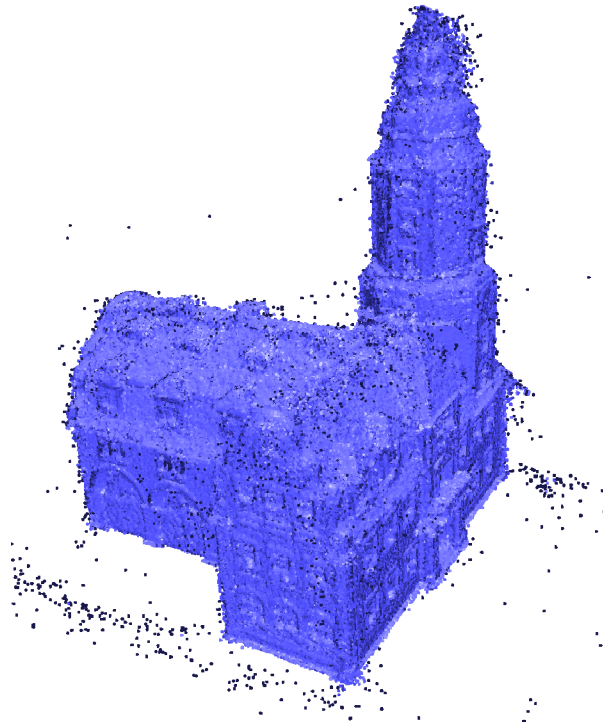


Abbildung 3: Per Laserscan aufgenommenes Objekt (Quelle: CGI, Uni Bonn)

2.3.3 3D-Modellierungsprogramme und Formate

Gemäß der in 2.2.2 zusammengefassten Entwicklung des CAD wird für die Erstellung und Bearbeitung von 3D-Architekturmodellen eine Vielzahl an unterschiedlichen 3D-CAD-, Modeler-, Animations-, Rendering- und Simulationsprogrammen verwendet, die auch aus branchenfremden Anwendungskontexten stammen können, beispielsweise aus der Automobil- oder Filmindustrie. Die Programme arbeiten mit heterogenen, bisweilen auch mit proprietären Dateiformaten. Eine Modellierungssoftware kann in der Regel aus ein und demselben Modell Ausgaben in mehreren Formaten und mit heterogenen qualitativen Attributen erzeugen. Einen Eindruck über die Verteilung der derzeit in der Architektur ange-

wendeten 3D-Formate für Modellierung und Austausch liefert die in Abschnitt 2.4.3 näher beleuchtete Umfrage unter Architekten. Wie in Abbildung 4 ersichtlich, ergibt sich ein heterogenes Bild aus Programmen zur Bearbeitung von Freiformflächen, klassischer orthogonaler Architektur, BIM-verarbeitenden oder rein geometrisch orientierten Planungswerkzeugen.

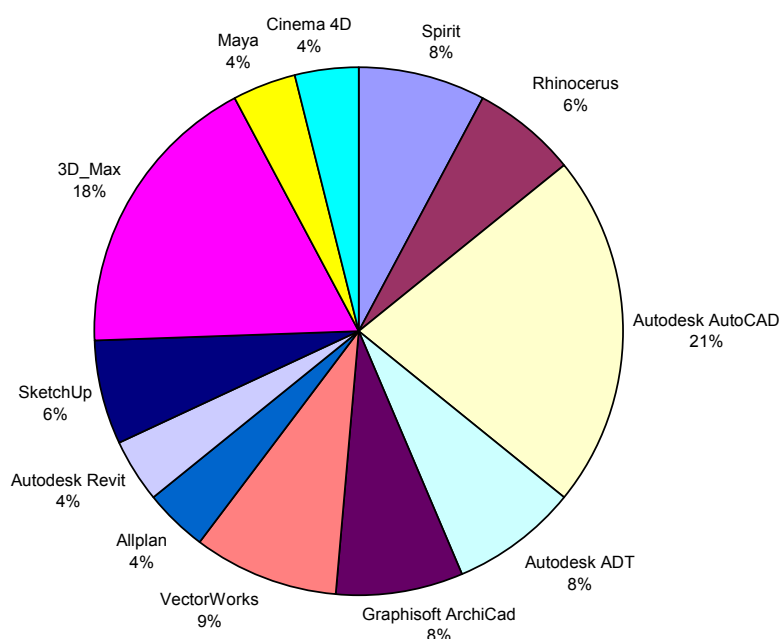


Abbildung 4: Verwendung von 3D-verarbeitenden Programmen (Auszug aus einer Befragung von Architekten, vgl. Abschnitt 2.4.3)

Aus der Bindung des Formats an seine Erstellungssoftware und aus dem geringen Interesse der Hersteller an einer kooperativen Formatentwicklung resultiert eine erschwerte Interoperabilität und eine hohe Abhängigkeit des Benutzers vom jeweiligen CAD-System, da insbesondere bei der Übernahme von Daten aus einem System ins nächste sowie generell bei Konvertierung bzw. Migration Informationsverluste auftreten können. Es existieren Bemühungen – ausgehend vor allem von der International Allianz for Interoperability (IAI) –, einheitliche Standards für die Beschreibung von 3D-Daten zu entwickeln, vgl. auch BIM / IFC im Abschnitt 2.2.2.

2.3.4 Modellaufbau

Im Hinblick darauf, ein Verständnis vom Aufbau architektonischer 3D-Modelle zu gewinnen, hilft die Perspektive auf die gedanklich-strukturellen Konzepte zur Problemlösung von architektonischen Planungsaufgaben im Hinblick auf die Techniken, Verfahren, Regeln und Prozesse, die für das Entwerfen konstitutiv sind. (Gethmann & Hauser, 2009) analysieren diese „Kulturtechnik Entwerfen“ im Zeitalter digitaler Technologien. Je nach Entwurfspraktiken können Modelle „bottom-up“, z.B. aus der Kombination von Einzelaspekten und Entwurfskomponenten, erstellt worden sein. Ein aus Einzelteilen konstruiertes Modell, z.B. ein Motorenmodell oder ein Gebäudemodell, wie in Abbildung 5 dargestellt, ist zu meist hierarchisch aufgebaut und enthält weitere z.T. komplexe Modelle. Diese Modelle können wiederum in sich geschachtelt sein, vgl. Ausführungen zu „Blöcken“ und parametrischem CAD in Abschnitt 2.2.2. Somit kann ein 3D-Modell, vor allem wenn es sich um ein konstruktionsorientiertes, präskriptives Modell handelt, eine Vielzahl an diskreten Einheiten enthalten, die einen eigenständigen Nachweis als 3D-Modell wünschenswert machen können. Hingegen liegt ein nach überwiegend axiomatisch „top-down“ modelliertes oder auch ein mittels Laserscan aufgenommenes und nach Tessellierung als Dreiecksnetz repräsentiertes Objekt in der Regel zunächst „als Ganzes“ vor, wird aber in anschließenden Prozessen zur (Binnen-) Strukturierung des Modells zumeist in semantisch unterscheidbare Einheiten aufgeteilt. (Thurrow, 2004, S. S.42) unterscheidet in einem Gebäudemodell zwei Arten von Primärstrukturen:

1. „eine Struktur von Bauteilen, die in ihrem konstruktiven Zusammenhang den Baukörper bilden“,
2. „eine Struktur von räumlichen Objekten, die zu Einheiten wie Raumgruppen, Etagen, Gebäudeteilen usw. gruppiert werden“.

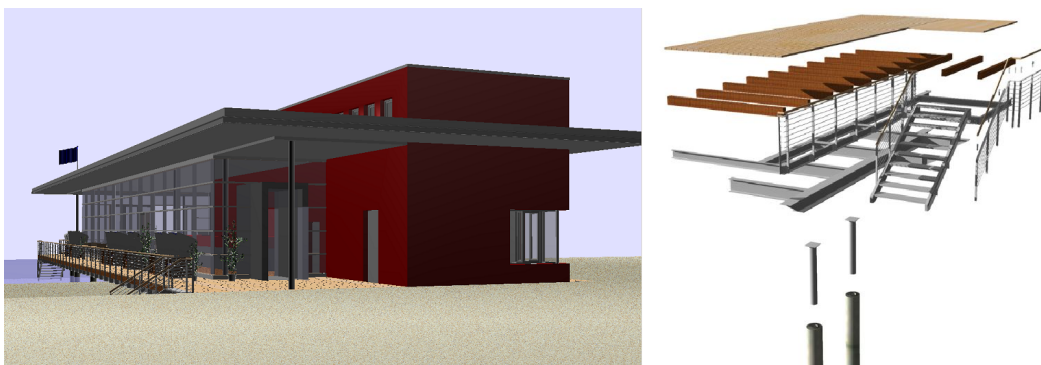


Abbildung 5: 3D-Gebäudemodell mit Visualisierung der im Gebäude enthaltenen strukturellen Komponenten (Quelle: Institut AIDA, LU Hannover)

Die oben genannte Strukturierung von 3D-Modellen kann als die zentrale Herausforderung für ein Metadatenmodell angesehen werden. Neben der Aufteilung des Modells auf verschiedene Dateien ist sie das wichtigste Merkmal der angesprochenen Komplexität, das 3D-Modelle von anderen Modellen unterscheidet. Im folgenden Teil wird näher auf die Rolle der im Modell enthaltenen Komponenten eingegangen.

2.4 Suche nach 3D-Modellen

Mit heterogenen Planungsaufgaben steigen Anzahl und Komplexität der Modelle, gleichzeitig benötigen Benutzer geeignete Werkzeuge zur Suche in 3D-Modellbeständen, um sie für eigene Entwürfe heranzuziehen. Um eine metadatenbasierte Beschreibung zu rechtfertigen, braucht es neben den aus implizit vorhandener Semantik resultierenden inhaltsbasierten Suchverfahren die Notwendigkeit, mittels Stichworten o.Ä. nach den annotierten Modellen suchen.

Im Internet existieren diverse 3D-Modell Repositorien mit teilweise entsprechend angepassten Suchschnittstellen. Typisches Beispiel ist das „3D Warehouse“ von Trimble (ehem. Google).²¹ Die meisten Modelle sind nur in ihrer spezifischen Suchumgebung such- und nutzbar. Einige sind frei verfügbar, andere benötigen

²¹ <http://sketchup.google.com/3dwarehouse/>

eine Registrierung zum (z.T. kostenpflichtigen) Download der Inhalte. In vielen Fällen werden die Inhalte direkt durch die Community (Architekten, Ingenieure, Designer, Studenten u.a.) erstellt. Aus Effizienz und Kostengründen werden 3D-Modelle in unterschiedlichen Kontexten adaptiert; und es ist davon auszugehen, dass die Nutzung dieser Bereitstellungs- und Austauschseiten für Modell-Content weiter zunehmen wird.

2.4.1 Suche im Entwurf: Case-based Reasoning

Das als Teil von sogenannten „Design-Methods“, vgl. (G. Schmitt, 1993), zur Formalisierung von Entscheidungsprozessen innerhalb komplexer Planungsaufgaben angewendete „Case-based Reasoning“ (fallbasiertes Schließen) ist eine in der Architektur lange bekannte Methode. Dabei werden architektonische Lösungen für ein bekanntes Problem gesucht und entsprechend der Lösung eines neuen Problems angepasst. Somit sind nur wenige Lösungen von Grund auf neu entwickelt.

Der Justiervorgang bzw. die Anpassung einer bereits bestehenden Architektur auf neue Probleme ist ein komplexer Vorgang. Die einfachste Stufe ist die direkte Übernahme. Auf der nächsthöheren Stufe werden Teile von bestehenden Architekturlösungen übertragen, andere werden geometrisch (z.B. in Bezug auf Materialien) angepasst. Im kompliziertesten Fall der Anpassung werden auch topologische, d.h. strukturelle Änderungen vorgenommen, vgl. (Blümel, Berndt, Ochmann, Vock, & Wessel, 2010).

2.4.2 Inspiration und Integration

Auf der Suche nach Informationen zu Planungsaufgaben suchen Architekten nach 3D-Modellen wie beispielsweise Gebäude, bauliche Details, Umgebungs- oder Stadtmodelle, technische Installationen, Möbel, Bauelemente oder andere Objekte zur Staffage von Visualisierungen oder Animationen. Zudem werden im wissenschaftlichen Kontext zunehmend 3D-Modelle erarbeitet, die für andere Wissenschaftler von Interesse sind.

Fall 1²²: Angenommen, ein Architekt beteiligt sich an einem Wettbewerb für einen Schulbau. Er erhält Unterlagen vom Auslober, z.B. das geforderte Raumprogramm und einen Plan des zu bebauenden Grundstücks. Der Architekt begeht das Grundstück, macht Fotos und gleicht den Plan ab. Um nicht die Umgebungsbebauung selbst eingeben zu müssen, sucht er nach bereits vorhandenen Städtebau- oder Umgebungsmodellen, in die er die neue Planung integrieren kann. Inspiration erhält er durch bereits geplante und ggf. gebaute Schulen, deren Modelle im Internet und in der Fachliteratur publiziert wurden, vgl. Case-based-Reasoning. Will er seine eigene Planung als 3D-Modell detailliert und perspektivisch illustrieren, benötigt er Bauelemente und Objekte zur Integration in seinen Entwurf, die allerdings aufwendig zu erzeugen sind. Im Fall von Bauelementen sind zudem semantische Informationen in den Modellen notwendig, beispielsweise zur Ausschreibung im weiteren Planungsverlauf. Aus diesen Gründen wird im Planungsprozess nach bereits in 3D modellierten Elementen und Objekten gesucht, die dann wiederverwendet werden können.

Fall 2: Ein Innenarchitekt und Tischler wird mit einem maßgeschneiderten Laden- ausbau beauftragt, dafür benötigt er Einbauobjekte für den Küchenbereich. Da das Erzeugen derselben aufwendig ist und im Fall von Bauelementen für die anstehende Ausschreibung auch Herstellerinformationen in diesen Modellen benötigt wird, sucht er auf den entsprechenden Herstellerseiten nach 3D-Modellen der zu integrierenden Komponenten.

Fall 3: Ein alternatives Szenario ist die Suche nach Gebäudemodellen für eine wissenschaftliche Arbeit. Angenommen, ein Doktorand schreibt an einer Arbeit über Grundrisstypen der Renaissance und sucht 3D-Rekonstruktionen der Villen Palladios. Die gefundenen Modelle werden von ihm nach Entwurfsmustern untersucht, er integriert die Modelldarstellungen in seine Arbeit und erzeugt in der Modellbauwerkstatt 3D-Drucke einzelner Gebäudemodelle. Er konstruiert ein eigenes, den möglichen Mustern entsprechendes parametrisches Modell einer fiktiven Villa Palladios.

²² Weiterentwicklung einer Darstellung aus (Blümel & Sens, 2009)

Die Beispiele zeigen eine Bandbreite an Verwendungsmustern, die sich in zwei Typen unterteilen lassen: 3D-Modelle können einerseits Inspirationsquelle, sowohl für eigene Entwürfe als auch für wissenschaftliche Arbeiten und Projekte sein, andererseits werden sie direkt in das eigene 3D-Modell integriert. Abbildung 6 illustriert diese beiden Anwendungstypen entlang des Prozesses der Planungsdetaillierung.

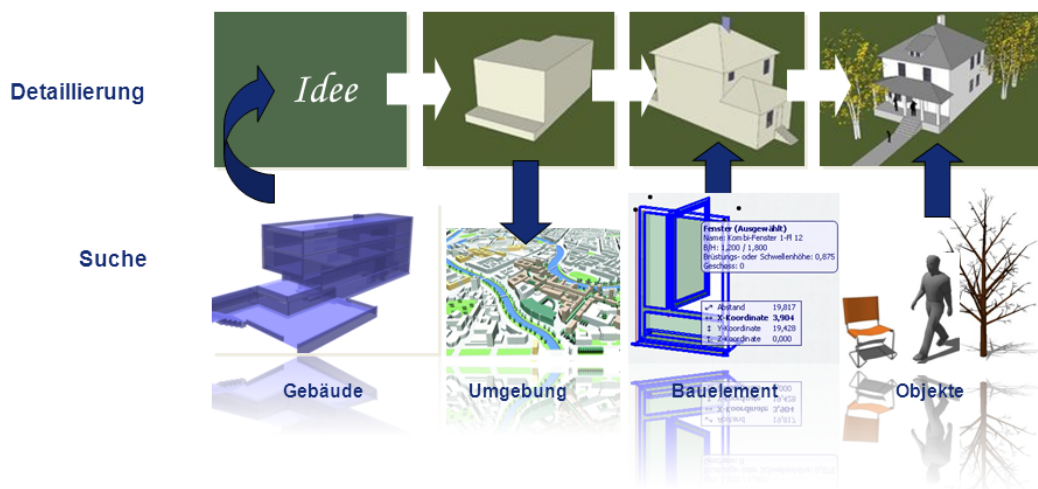


Abbildung 6: Verwendung von 3D-Modellen entlang des Prozesses der Planungsdetaillierung

2.4.3 Befragungsergebnisse

An die Such- und Verwendungsszenarien von architektonischen 3D-Modellen aus Abschnitt 2.4.2 anknüpfend werden an dieser Stelle die Ergebnisse einer Befragung²³ von 61 Architekten, Universitätsmitarbeitern und Studenten hinzugezogen, um die für eine 3D-Modellsuche notwendigen Metadaten zu konkretisieren.

²³ Die Befragung wurde von der Autorin im Kontext der PROBADO-Projektstätigkeit, vgl. (Berndt, Blümel, Clausen, u. a., 2010), an der TIB Hannover mit Unterstützung der Architektenkammer Niedersachsen ab 2006 initiiert. In dieser Arbeit werden, auf erste in (Blümel & Sens, 2009) dargestellte Ergebnisse aufbauend, erheblich mehr und ausschließlich komplett abgeschlossene Umfragen berücksichtigt. Der Modelltyp „3D-Detail“ ist in die Darstellung in dieser Arbeit nicht integriert.

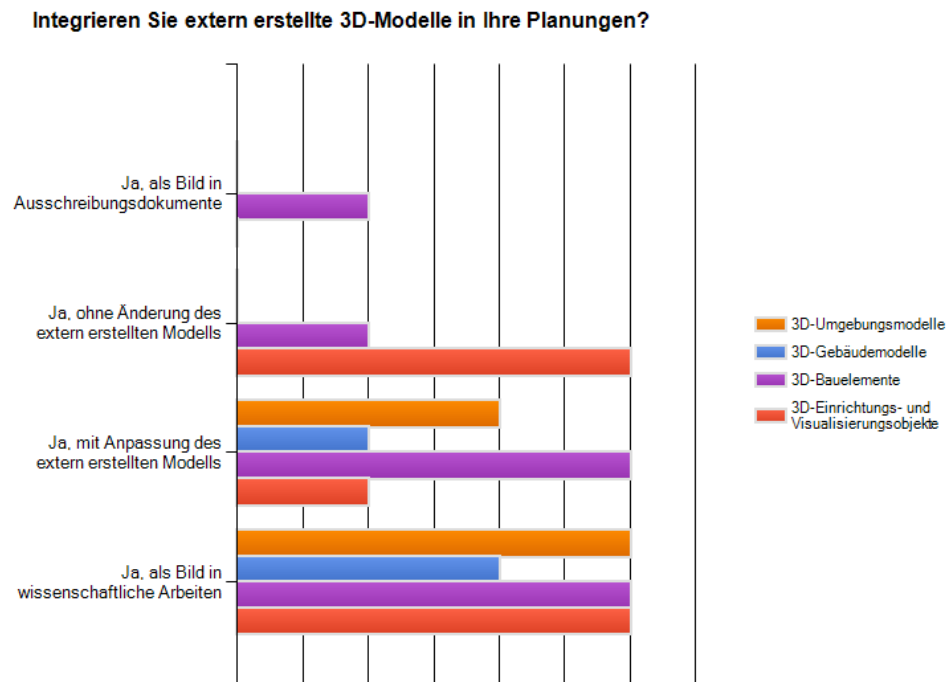


Abbildung 7: Verwendungsszenarien extern erstellter 3D-Modelle

Die Befragung liefert analog zu den vorgehenden Suchszenarien zunächst Hinweise über den Verwendungstyp, also Inspiration oder Integration, vgl. Abbildung 7 und, speziell hinsichtlich Planungsphasen nach HOAI, Abbildung 8. Gebäudemodelle werden überwiegend zur Inspiration verwendet, Umgebungsmodelle, Bauelemente und Objekte zur Integration in den eigenen Entwurf, siehe auch Darstellung in Abbildung 6. Weiterhin lässt sich aus den Ergebnissen schließen, dass die selbst planenden Nutzer von 3D-Inhalten schon relativ früh im Planungsprozess weniger an Modellen ganzer Gebäude interessiert sind als vielmehr an den Objekten, Bauelementen und Details innerhalb eines Gebäudemodells, vgl. Abschnitt 2.3.4.

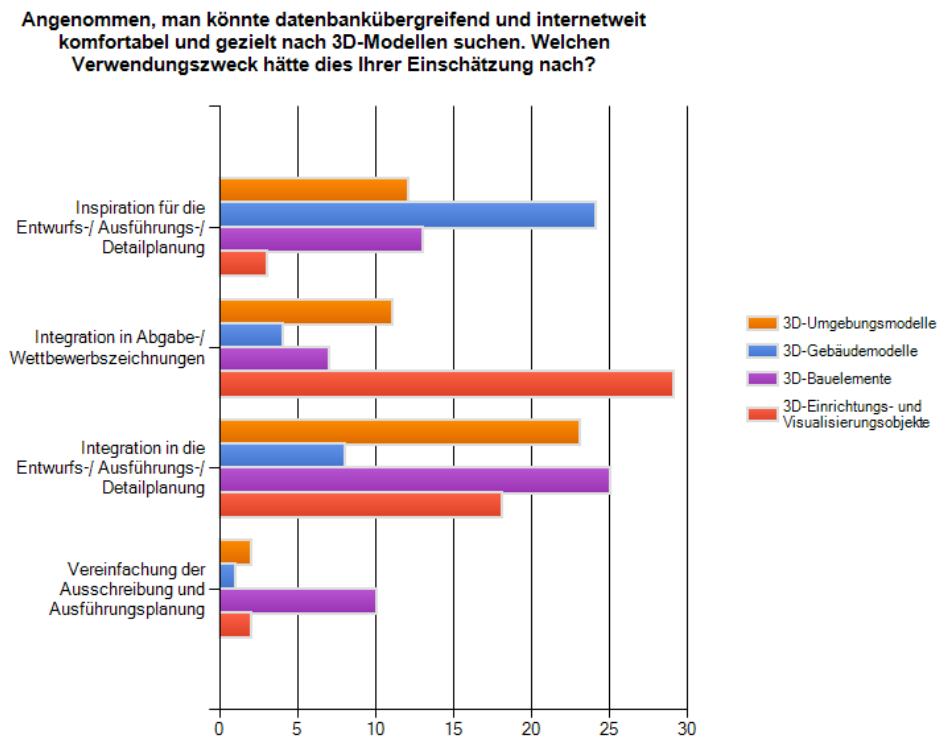


Abbildung 8: Verwendung von 3D-Modellen in der Planungspraxis

Ein klares Signal für die Notwendigkeit von Metadaten für 3D-Modelle gibt das nächste Befragungsergebnis. Die befragten Architekten geben den klassischen textuellen Suchmöglichkeiten (an erster Stelle der stichwortbasierten und an zweiter Stelle der klassifikationsbasierten Suche) den Vorzug gegenüber zeichnerischen und anderen inhaltsbasierten Recherchemöglichkeiten, und zwar über alle Modelltypen hinweg. An dritter Stelle wird die Anfrage mit einem Beispielmmodell²⁴ genannt, vgl. Abbildung 9.

²⁴ sogenanntes „Query-by-Example“

Welche Suchmöglichkeiten sollten für die jeweiligen Modelltypen angezeigt werden?

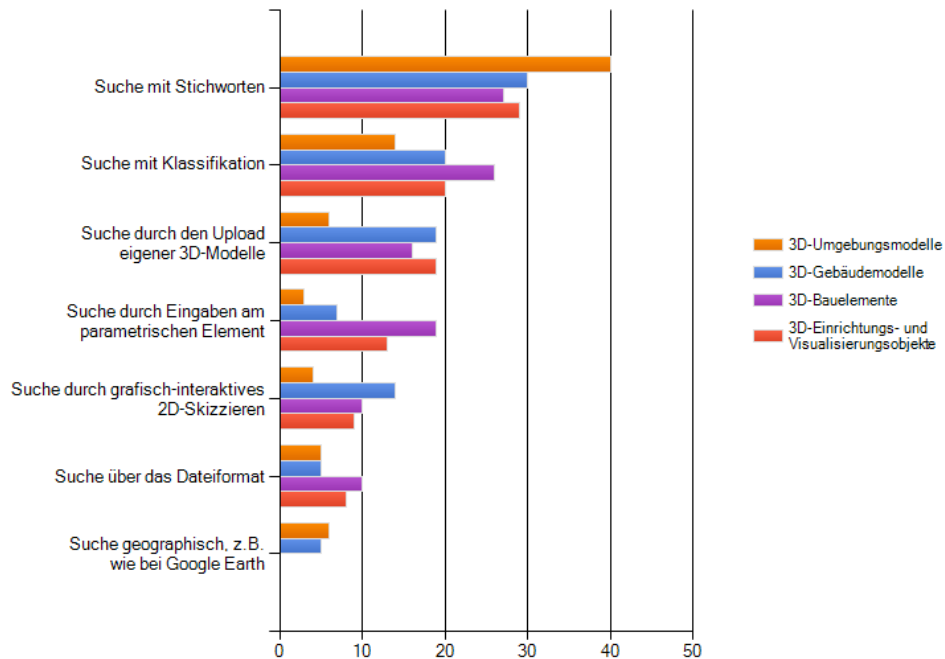


Abbildung 9: Gewünschte Suchmöglichkeit / 3D-Modelltyp

Gleichwohl ist unter bestimmten Voraussetzungen eine nicht-textbasierte Suche vorzuziehen. Die Befragungsergebnisse sollen daher nicht generell die metadatenbasierte Annotation und Suche rechtfertigen, auch vor dem Hintergrund dessen, dass die Probanden die Möglichkeiten einer inhaltsbasierten Suche mangels Erfahrung in der Benutzung wahrscheinlich schwer abschätzen konnten. Vielmehr ist durch den anfänglichen Mangel an separierten informationstragenden Einheiten bei 3D-Modellen wie auch anderen nichttextuellen Materialien eine Annotation mit Metadaten notwendig. Ein Vorteil der Metadaten ist die explizite Beschreibung und damit Vernetzbarkeit mit weiteren Medienformaten, die ähnliche Informationseinheiten beinhalten.

Optimale Suchmöglichkeiten nach 3D-Modellen sollten daher eine Kombination aus textbasierten und inhaltsbasierten Anfragemöglichkeiten darstellen, vgl. (Bernd, Blümel, Krottmaier, Wessel, & Schreck, 2009). So könnte eine Suche mit

Text oder Query by Example starten und dann z.B. mittels Filterung der Suchergebnisse (Facettierung) weiter verfeinert werden.

2.5 Anforderungen an Metadaten für 3D-Modelle

Anforderungen hinsichtlich technisch-struktureller Aspekte

Zunächst wird eine Auswahl an typischen Attributen auf Dateiebene mit beispielhaften, einer Suche vorangehenden Fragen angeführt:

Format: Passt die Datei zum CAD-Programm?

Anzahl der Polygone: Wie detailliert²⁵ ist das Modell?

Dateigröße: Wie groß ist der Umfang des Downloads?

Außerdem gibt es Parameter, welche in speziellen Zusammenhängen erforderlich sind, z.B. für Ausschreibungen oder Bebauungspläne. Beispiele solcher Attribute, die wie die vorgenannten Attribute ebenfalls quantifizierbar sind, wären:

Größe der Bruttogeschossfläche, Anzahl der Stockwerke, Traufhöhe oder Dachneigung.

Anforderungen hinsichtlich inhaltlich-kontextueller Aspekte

Um 3D-Modelle in ihrem inhaltlichen Kontext zu beschreiben, gilt es zunächst, die wichtigsten Konzepte zu identifizieren. Die Attribute sind unabhängig davon, um welchen konkreten Dateityp es sich handelt. Wie anhand der Fallbeispiele in 2.4.2 eingeführt, könnten Anwender nach 3D-Modellen anhand folgender Kriterien suchen:

- Objekte: Welches konkrete Gebäude oder Objekt stellt das 3D-Modell dar?
- Orte: Wo findet der Wettbewerb statt? / Wo ist das Haus zu bauen? / o.ä.
- Contributoren: Wer ist der Modellersteller? / Wer ist der Architekt? / Wer hat das Seminar initiiert? / o.ä.

²⁵ Die Detaillierung lässt auf die Planungsphase, beispielsweise Entwurfsplanung oder Ausführungsplanung, schließen und in Konsequenz darauf, ob das Modell zur Inspiration oder eher zur Integration zu gebrauchen ist.

- Konzepte: Welcher Baustil liegt dem Gebäude zugrunde? / Welches (eindeutig benennbare)statische System wurde angewendet? / Welche Form oder Funktion hat das gesuchte Bauteil? Anzumerken ist, dass an die Übergänge zur inhaltsbasierten Suche fließend sind. Auch mittels Skizziersuche oder Anfrage mit einem 3D-Beispiel (Query-by-Example) kann und sollte nach Formen und Strukturen gesucht werden, vgl. (Bernd u. a., 2009), zumal die textbasierte Ausdrucksstärke bei solchen Konzepten tendentiell ungenau wird.
- Ereignisse: Welcher Wettbewerb / welches Seminar / o.ä. war Anlass zur Erstellung des 3D-Modells?

Zusammenfassung

Unter Berücksichtigung der in diesem Kapitel eingeführten inhaltlichen und technischen Vorgaben ist bei der Entwicklung eines neuen Metadatenmodells also Folgendes zu beachten:

- Je nach Domäne und Anwendungsbereich muss die Diversifizierung von Typen und Formaten sowie geometrischen Modellrepräsentationen berücksichtigt werden,
- Eine evtl. vorhandene Binnenstruktur bzw. inhärente Entitäten müssen abgebildet werden können,
- Dokument-Aggregationen müssen unterstützt werden können,
- Quantifizierbare domänenspezifische Attribute müssen flexibel integriert werden können,
- Der Modellkontext muss abgebildet / verbunden werden können.

3 Informationsextraktion und Kontextualisierung

Information Retrieval²⁶ Systeme (IR-Systeme) dienen zum Auffinden von Informationsobjekten, die inhaltliche Übereinstimmung mit / inhaltliche Relevanz zu der Suchanfrage aufweisen. Während das reine Daten-Retrieval ohne die semantische Interpretation der im System enthaltenen Objekte hinsichtlich implizit vorhandener Informationen stattfinden kann, ist eine solche Interpretation beim Informations-Retrieval zwingend notwendig. So möchte man z.B. alle 3D-Modelle finden, die einen Tisch darstellen, d.h. anhand ihrer semantischen Interpretation dem Konzept Tisch zugeordnet werden können.

Wie in Abschnitt 2.4.3 anhand der Befragungsergebnisse erörtert wurde, werden architektonische 3D-Modelle insbesondere anhand ihrer textuellen Annotation gesucht. Selten ist der o.g. Tisch aber auch als solcher annotiert. Dieses Kapitel behandelt die Anreicherung von 3D-Modellen mit inhaltlichen Annotationen: die inhaltsbasierte Erschließung.

3.1 Retrieval

Meistens kennt der Benutzer weder das gesuchte Informationsobjekt noch dessen genaue Bezeichnung, und oft weiß er nicht, wo es zu finden ist und wie die nötige Suchanfrage für das sogenannte „known-item“ exakt formuliert wird. Ähnlich gestaltet es sich, wenn nicht nur ein Objekt sondern eine größere Ergebnismenge gewünscht wird und erschöpfende Informationen zu einem mittels Suchanfrage exakt bestimmbar Thema gesucht werden. Im Gegensatz dazu ist die explorative Suche eher ungerichtet, d.h., der Benutzer hat keine exakte Vorstellung von der Ergebnismenge, oder er weiß zwar ungefähr, wonach er sucht, kann die Anfrage jedoch nicht exakt formulieren. Beispielhaft genannt seien hier das „Berry-Picking Model“ einer sich iterativ verfeinernden Suchanfrage, vgl. (Bates, 1989),

²⁶ Information Retrieval wird sowohl der Informationswissenschaft als auch der Informatik und Computerlinguistik zugeordnet und beschäftigt sich mit IT-unterstützter Suche nach komplexen Inhalten.

sowie die von einem oder mehreren besonders passenden Informationsobjekten ausgehende „(Citation) Pearl Growing“-Strategie, vgl. (Efthimiadis, 1996, S. 127ff). Es sei hier erneut angemerkt, dass Exploration und Browsing für 3D-Modelle nicht nur mit textuellen Annotationen möglich ist, vgl. Beispiele in (Bernd u. a., 2009).

Der Benutzer kann die Art der Informationen, die er sucht, selten in präzise und zielführende Suchbegriffe fassen. Wird die Suchanfrage als Text in natürlicher Sprache behandelt, wird gemäß einer Volltextsuche nur nach deskriptiven Schlüsselworten gesucht, und die Ergebnismenge ist relativ gering. Diese Methode hat zudem den großen Nachteil, dass die Person, die sucht, nicht automatisch dieselben Begriffe benutzt wie die Person, die klassifiziert hat.

Im Gegensatz dazu arbeitet die semantische Suche auf der Grundlage semantischer Entitäten. Der durchsuchbare Bestand wird entsprechend aufbereitet, und in der Suchanfrage werden eindeutige semantische Entitäten verwendet, sodass die erzielten Suchergebnisse eine höhere Genauigkeit und Vollständigkeit aufweisen. Darüber hinaus können assoziative Beziehungen zwischen den semantischen Entitäten genutzt werden, z.B. für die o.g. explorative Suche, für Browsing oder für Suchvorschläge.

Um also die o.g. Suchstrategien, insbesondere die explorative Suche, zu ermöglichen und dem Benutzer die gewünschte Ergebnismenge an 3D-Modellen zu liefern, ist eine systematisierte Aufbereitung und Verfügbarmachung von sowohl modellinhärenten als auch den Modellkontext betreffenden Informationen als textuell durchsuchbare und semantisch unterscheidbare Entitäten notwendig. Im Folgenden geht es um die Bereitstellung geeigneter Metadaten, die ein effektives Retrieval relevanter Modelle ermöglicht.

3.2 Wissen organisieren

3.2.1 Daten – Information – Wissen

Zum Verständnis der Methoden, mit denen eine relevante Information aus einem Informationsobjekt extrahiert und Wissen organisiert werden kann, ist eine Ab-

grenzung der Begriffe Daten, Information und Wissen notwendig. (Rehäuser & Krcmar, 1996, S. 7) setzen die Begriffe innerhalb einer definierten Hierarchie semiotisch in Beziehung, siehe Abbildung 10.

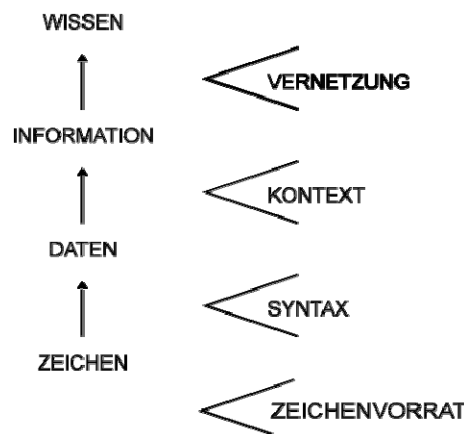


Abbildung 10: Beziehung zwischen den Begriffen Wissen, Information, Daten und Zeichen nach (Rehäuser & Krcmar, 1996)

Daten sind das Medium, mit dem Informationen abgelegt und transportiert werden. Sie bestehen aus verabredeten und bekannten *Zeichen*, entsprechen also einer bestimmten Syntax. Aus Daten werden *Informationen*, wenn sie in einen sinngebenden Zusammenhang gestellt und damit um eine semantische Dimension erweitert werden. Information stellt eine unter Berücksichtigung von Syntax und Semantik entstandene Mitteilung dar.

Auf diese Weise zustande gekommene Informationen bilden die Basis von *Wissen*. Erst eine Vernetzung von Informationen ergibt Wissen, demnach ist die Transformation von Informationen in Wissen von einer kontextualisierten Sicht auf die Informationsobjekte geprägt. Diese Sicht ist das Ergebnis der Verarbeitung der Informationen durch das menschliche Bewusstsein.

Analog ist das bereits erwähnte „Semantic“ Web ein von einer kontextualisierten Sicht auf Informationsobjekte geprägtes Paradigma für den Umgang mit ‚Information‘ und deren Transformation in ‚Wissen‘, vgl. Diskussion in Abschnitt 3.5. (Stachowiak, 1973) modelliert Wissen als Abstraktion der Realität entsprechend

der subjektiven Wahrnehmung des Modellkonstruktors, vgl. Kapitel 1. Die formalisierte Darstellung lässt sich wiederum für weitere wissensbildende Prozesse verfügbar machen, sodass Wissen als kontextualisierte Information und Grundlage für neues Wissen betrachtet werden kann. In ähnlicher Weise definiert (Aune, 1970): „Wissen ist Information in Kontext.“

3.2.2 Erschließung (traditionell)

Wissensorganisation beschreibt „Methoden und Systeme zur Erschließung und Organisation von Informationen“ bzw. von „dem mit ihnen ausgedrückten Wissen“.²⁷ Im Fokus von Wissensorganisation steht das Information Retrieval, d.h., die Wissensorganisation soll den Retrieval-Prozess maßgeblich verbessern. Metadaten spielen für die / bei der Erschließung und Strukturierung bestehender Wissensbestände eine richtungsweisende Rolle.

Um Wissen erschließbar zu machen, werden zunächst Methoden und Werkzeuge angewandt, die Informationsobjekte durch ein Stellvertreterobjekt, das Surrogat, abbilden. Das Surrogat enthält verdichtete formale und inhaltliche Angaben zu den vorliegenden Informationen über das jeweilige Objekt, es enthält also eine Zusammenfassung oder auch eine Abstraktion der dem jeweiligen Objekt inhärenten Informationen. So sind im Surrogat zwar möglichst viele Informationen über das Original enthalten, aber so aufbereitet und komprimiert, dass das spätere Information-Retrieval erleichtert wird. Dabei ist die geringere Menge an Daten, die während der Suche verarbeitet werden muss, von Vorteil, denn in der Regel sind Surrogate erheblich kleiner als die Objekte, auf die sie verweisen. Im Abschnitt 3.5 wird auf Wissensorganisation und formalisierte Beschreibungen im LOD-Kontext erneut eingegangen.

²⁷ vgl. Definitionen in <http://de.wikipedia.org/wiki/Wissensorganisation> sowie verbundene Links

Traditionell wird zwischen Formalerschließung²⁸ und Inhalterschließung²⁹ unterschieden, d.h. zwischen formalen Merkmalen (z.B. Titel, Autor, das Datum der Veröffentlichung) und inhaltlichen Merkmalen (zur Erfassung der Begriffs- und Bedeutungs-Systematik, z.B. behandelte Gegenstände, Personen, geografische und zeitliche Bezüge des Themas), vgl. (Umlauf, 2000, Kapitel 1.4, 1.5). Ein Informationsobjekt kann gleichzeitig mit mehreren Verfahren erschlossen werden³⁰. Hierzu werden vornehmlich bibliothekarische Normdaten verwendet, also kontrollierte Begriffe aus Wissensordnungen wie Ontologien, Thesauri, Klassifikationen zur Beschreibung von Ressourcen, wobei zwischen klassifikatorischer und verbaler Inhalterschließung unterschieden wird. Ein Beispiel für das Anwendungsgebiet Architektur ist der Getty Art & Architecture Thesaurus³¹. (Stock & Stock, 2008, S. 42) stellen fest, dass die Ausdrucksstärke der Verfahren, d.h. die Anzahl der Relationen zwischen den Begriffen des jeweiligen Systems, mit zunehmendem Umfang der Wissensdomäne abnimmt. Aktuell wird ein Einbeziehen von weniger regelgeleiteten Methoden der Wissensorganisation diskutiert und Inhalterschließung als offen für Verfahren wie Folksonomies oder Tagging definiert, siehe beispielsweise (Eberhardt, 2011).

3.2.3 Records und „Datensilos“

In einem solchen Ordnungssystem spielen Katalogeinträge eine zentrale Rolle, und so ist in der Bibliothekspraxis die klassische Form eines Metadatensatzes / Metadaten-„Records“ das Katalogisat³². Katalogbasierte (digitale) Bibliotheken vermitteln Zugang zu Informationsobjekten, d.h. sie vermitteln einen „Zeiger“

²⁸ auch als *Katalogisierung* bezeichnet, umfassende Einführung z.B. in (Wiesenmüller, Heidrun, 2004)

²⁹ auch als *Sacherschließung* bezeichnet, umfassende Einführung z.B. in (Bertram, 2005)

³⁰ durch Metatexte oder strukturierte Beschreibungen der Informationsobjekt-Inhalte

³¹ <http://www.getty.edu/research/tools/vocabularies/aat/index.html>

³² definiert analog (Kent, 1979) als “[...] a fixed sequence of field values, conforming to a static description usually contained in catalogs and/or in programs.”

von den Metadaten auf die Objekte, meist Teil des eigenen Bibliotheksbestandes bzw. Kataloges³³.

Kataloge bieten eine Übersicht zu in sich abgeschlossenen bzw. abschließbaren Sammlungen, deren verzeichnete Objekte aufzählbar sind und durch ihre Metadaten in ein Verhältnis zu anderen Objekten im treten (z.B. Suche aller Objekte desselben Autors, die mit demselben Schlagwort ausgezeichnet sind). Der Bearbeitungsprozess (Einarbeitung, Speicherung, Beschreibung und Bereitstellung) fokussiert auf Objekte als Informationscontainer, weniger auf den Inhalt dieser Container. Herausforderungen dieser Bearbeitungslogik in Hinblick auf das zur Verfügung Stellen kontextualisierter Information wurde bereits diskutiert und wird in Abschnitt 3.5 vertieft. Karen Coyle schreibt 2010

“With the expansion of the Web, online catalogues have a new context that involves both opportunities and challenges. The opportunities are related to the effective infrastructure for sharing and dissemination. Among other things, the challenges relate to the existing library standards for document description – metadata. These standards were made in a different technological era.”³⁴

Im Folgenden werden Metadaten auch unter dieser Prämisse diskutiert. Das Kapitel dient zunächst einer aktualisierten Definition des Metadatenbegriffs sowie einer Übersicht zum Gebrauch von Metadaten. Es folgt eine kurze Erläuterung der Attribute sowie die Einführung der im Kontext dieser Arbeit für das Konzept Metadaten wichtigen Begriffe.

³³ „Katalog“ definiert sich lt. Duden als ein „nach einem bestimmten System geordnetes Verzeichnis von Gegenständen, Namen o. Ä.“ Vgl. auch <http://gordondunsire.com/presentations.htm> A short history of the evolution of the catalogue record.

³⁴ in <http://journal.code4lib.org/articles/6424>

3.3 Metadaten: Beschreibende Funktion

Metadaten – routinemäßig als „Daten über Daten“ definiert – sind strukturierte Daten, welche (auf einer höheren, sogenannten *Meta*³⁵-Ebene) die Merkmale und Inhalte von Referenzobjekten strukturiert beschreiben. Einen praxisnahen Überblick zu Metadaten gibt beispielsweise (Caplan, 2003). Metadaten sind immer relativ, d.h. Metadaten sind Daten über beziehungsweise zu anderen Daten. Die Metaebene kann ihrerseits auch als Objektebene betrachtet werden, somit können auch Metadaten zu Metadaten gebildet werden. Der Begriff Metadaten ist nicht neu und schon lange in Bibliotheken gebräuchlich. So beschrieb ein Metadaten-satz für eine sehr lange Zeit das Katalogisat³⁶, und damit einhergehend beschrieb die Erstellung von Metadaten das mit traditioneller Katalogisierung gleichzusetzende Konzept der Erschließung bzw. Beschreibung von Informationsobjekten, vgl. Abschnitt 3.2.2. Eine solche Sicht wird durch Definitionen wie die folgende in (Heery, 1996) untermauert:

“The familiar library catalogue record could be described as metadata in that the catalogue record is 'data about data'.”

Die normierten Katalogisierungsregeln der Bibliothekscommunity für den überschaubaren eigenen Katalog sind jedoch nur schwerlich auf die große, ungeordnete Menge von Ressourcen in elektronischen Repositorien wie dem WWW zu übertragen. Die Neudefinition des Konzepts Metadaten umfasst daher auch die Suche nach neuen Möglichkeiten, Ressourcen und Informationen optimiert für einen effektiven Einsatz in diesen Suchräumen zu beschreiben.

3.3.1 Aktualisierte begriffliche Annäherung

Wie im Kapitel 1 erörtert, haben sich durch das WWW eine gewisse Anzahl grundlegender Parameter unseres Umgangs mit Informationsobjekten verändert

³⁵ Der griechische Ausdruck *Meta* bezeichnet eine höhere bzw. fundamentalere Ordnung.

³⁶ Ein Metadaten-Datensatz besteht hier aus einer Reihe von vordefinierten Datenrepräsentanten, die Attribute einer Ressource repräsentieren, wobei jedes Element einen oder mehrere Werte eines definierten Wertebereichs haben kann.

und die Definition des Dokumentenbegriffs fundamental gewandelt. Die Annäherung an den aktuell in verschiedenen Kontexten (insbesondere WWW und Linked Data, zunehmend in Überschneidung mit digitalen Bibliotheken) sehr aktiv verwendeten Begriff Metadaten impliziert die Notwendigkeit, das Verhältnis zwischen Metadaten und referenzierten Elementen zu überdenken. Speziell hinsichtlich elektronisch vorliegender Informationsobjekte ist diese Annäherung – ausgehend vom deskriptiven Katalogparadigma wie in Abschnitt 3.2 beschrieben – eine grundlegende Voraussetzung für eine Neudefinition der Rolle von Bibliotheken als Informationsdienstleister.

Anhand der folgenden Metadaten-Definition von (Berners-Lee, 1997) kann die signifikante Aufweitung derselben über den katalogbasierten Anwendungsbereich Bibliothek deutlich wahrgenommen werden. So seien Metadaten

“Machine understandable information about web resources or other things”

und unterscheiden sich damit als Garant zuverlässiger Meta-Beschreibung für die Authentifizierung einzelner Informationsobjekte zwar nicht substantiell vom Katalogisat, jedoch treten an die Stelle der Nutzer der Information nun Maschinen, womit sich der Anwendungskontext verschoben hat. So wird außerdem klar von maschinen-verstehbar und nicht nur von maschinen-lesbar gesprochen, womit die Maschinen sogar in die Lage versetzt werden sollen, Inferenzen³⁷ zu bilden, ähnlich der menschlichen Möglichkeit, aus kontextualisierter Information Wissen abzuleiten, vgl. Abschnitt 3.2.1.

Im weiteren Verlauf der o.g. Quelle gibt Berners-Lee mit

“In the future, when the metadata languages and engines are more developed, it should also form a strong basis for a web of machine understandable information about anything: about the people, things, concepts and

³⁷ Auf den Ausblick auf das Semantic Web als hermeneutisches Web und die Neudefinition durch Berners-Lee (Syntactic Web) soll an späterer Stelle erneut eingegangen werden.

ideas. We keep this fact in our minds in the design, even though the first step is to make a system for information about information.”

auch einen Ausblick auf die Entitäten, welche in der Lage sind, die Information aus dem Objekt selbst sowie aus dessen Kontext zu beschreiben. Der Katalog mit den einzelnen Referenzobjekt-Beschreibungen, den Katalogisaten, weitet sich damit auf – zu einem System von Entitäten als (vernetzte) Informationen über Informationsobjekte. Dabei ist nicht zwingend gegeben, dass nur professionelle Katalogisierer die Urheber der Informationsbeschreibungen sind, die Provenienz dieser Daten ist damit weniger eindeutig.

Die beschriebene Annäherung an den Begriff Metadaten, resultierend aus einer WWW-originären Informationsinfrastruktur, deckt sich nur noch teilweise mit dem deskriptiven Katalogparadigma.

3.3.2 Anwendung

(Coyle, 2004) definiert mit

“Metadata is...constructed...for a purpose...to facilitate an activity...”

Metadaten als zweck- und anwendungsbezogene Daten. Beispielsweise ist der Zweck das Finden und Abrufen von Information(subjekten) zu einem bestimmten Thema innerhalb eines definierten Suchraums. Wie in Abschnitt 1.2.1 erörtert, steigt das Volumen an digital verfügbarer Information, was bei einer Suche im Allgemeinen eine große Ergebnismenge, auch als „information overload“ bezeichnet, mit einer niedrigen Rate an präzisen Ergebnissen produziert (high Recall, low Precision³⁸). Dies erschwert dem Benutzer, innerhalb der großen Fülle an Daten zwischen den für ihn relevanten und irrelevanten Informationen zu unterscheiden.

Metadaten sollen als systematische Methode zur Beschreibung von Ressourcen und damit die Verbesserung des digitalen als auch intellektuellen Zugangs zu ih-

³⁸ Kennzahlen des Information Retrievals: Recall (Prozentsatz der relevanten Dokumente / Objekte, die gefunden werden) und Precision (der Prozentsatz der gefundenen Dokumente / Objekte die relevant sind).

nen dabei helfen, exaktere Suchergebnisse zu erzielen. So spezifizieren (Dempsey & Heery, 1998):

“...metadata is data associated with objects which relieves their potential users of having to have full advance knowledge of their existence or characteristics. A user might be a program or a person.”

Metadaten sollen Information also in der Weise sichtbar machen, dass sie ermöglichen zu entscheiden, ob der Abruf eines Informationsobjekts lohnt oder nicht. Wenn eine Ressource die Bereitstellung lohnt, dann lohnt es demnach auch, sie mit Metadaten zu beschreiben damit sie besser gefunden wird. Metadaten bieten damit eine wesentliche Verbindung zwischen dem Ersteller und dem Benutzer von Information.

Neben der beschreibenden Funktion ist die digitale Struktur ein wichtiger Aspekt des Konzepts Metadaten.

3.3.3 Struktur / Elementtypen

Metadaten werden üblicherweise wie folgt typisiert:

Beschreibung / Innenseite

Deskriptive Metadaten stellen Information über den Inhalt eines Referenzobjekts sowie dessen inhaltlichen Kontext bereit. Die Erzeugung von deskriptiven Metadaten, also die Formal- bzw. Inhaltserschließung, wurde in Abschnitt 3.2.2 dargelegt.

Verwaltung / Außenseite

Administrative Metadaten werden zur Verwaltung von Ressourcen benutzt und beinhalten u.a. Informationen zur Provenienz von Metadaten, deren Nutzungsbedingungen sowie Angaben zur technischen Verarbeitung und Archivierung.

Technische Charakteristika von Referenzobjekten, z.B. Informationen über Formate, werden den Typen je nach Anwendungszusammenhang den beschreibenden oder den administrativen Metadaten zugerechnet. Informationsobjekte besitzen in

der Regel zudem ein Metadaten-Element, das eine Identifikationsnummer (ID) als eindeutige Kennung beinhaltet.

Struktur / Container

Soweit es sich um ein mehrteiliges Informationsobjekt handelt, beschreiben *strukturelle Metadaten* dessen Datenstruktur, also die Logik oder die Beziehungen zwischen den Teilen. Relationen zu anderen Objekten können ebenfalls als Teil der Strukturdaten angesehen werden.

Metadatenelemente, die sich auf ein Referenzobjekt beziehen, werden zu einem Metadatensatz, vgl. 3.2.3, bzw. einer Dokumentationseinheit zusammengefasst, wobei Metadaten-Containerformate auch mehrere Informationsobjekte innerhalb einer Dokumentationseinheit beschreiben können. (Umlauf & Gradmann, 2011) definieren ferner zu strukturellen Metadaten:

„Je nach Granularität kann diese Beschreibung von einem einfachen Inhaltsverzeichnis bis zur detaillierten Repräsentation der Binnenstruktur reichen, sodass hier die Grenze zwischen Metadaten und Objektdaten fließend ist.“

Dieser nicht eindeutig abzugrenzende Übergang lässt sich insbesondere bei den neueren unter BIM zusammengefassten 3D-Modellbeschreibungen feststellen, vgl. Abschnitt 2.2.2.

Metadaten kommt daher im Kontext aggregierter und vernetzter Informationsobjekte mit nicht eindeutig beschreibbaren Grenzen des eigentlichen digitalen Dokuments, vgl. Abschnitt 1.2.4, einerseits eine dokument-konstituierende Aufgabe zu, andererseits können sie eben diese Vernetzung abbilden und ermöglichen. Weiterhin können sie in der Lage sein, objekt-inhärente Granularität abzubilden.

3.3.4 Ebenen

Metadaten repräsentieren einzelne Ressourcen mittels strukturierter Beschreibungen. Um Metadaten zu verstehen, muss deren jeweilige Struktur bekannt sein – manifestiert als spezifischer Metadaten-Standard und ausgedrückt in Schemata, Datenmodellen, Elementen, etc. Standardisierte Metadaten vereinfachen den Ver-

gleich von ähnlichen Elementen, beschrieben durch ähnliche Begriffe und Konstrukte im Gegensatz zu Beschreibungen, die undefinierter Form³⁹, beispielsweise als Freitext vorliegen. Dabei lassen sich zwei Ebenen spezifizieren:

1. die Informationseinheiten selbst, also das Vokabular, welches aus einem definierten Satz von Metadatenelementen besteht, vgl. 3.3.3, die verwendet werden, um Informationsobjekte zu beschreiben;
2. die Anordnung der oben genannten Informationseinheiten, also das Datenmodell.

Demnach bestehen die ein Referenzobjekt beschreibenden Metadaten jeweils aus den verschiedenen Metadatenelementen, die das Vokabular definieren und dazu verwendet werden, den Inhalt eines Metadatenschemas auszudrücken. Die Elemente werden anhand des Metadaten-Datenmodells organisiert, d.h. es erstellt die Regeln, nach denen die Elemente strukturiert werden. Zur Formulierung von Modellen gibt es unterschiedliche Methoden und Sprachen, beispielsweise das Entity Relationship Modell (ERM) oder die Unified Modeling Language (UML)⁴⁰. Methoden zur Modellbeschreibung definieren die innerhalb eines Modells verwendeten Terme (classes, properties, data types), spezifizieren die Beziehungen zwischen Ressourcen / Objekten und geben Überblick über die Informationsstrukturen. Das Datenmodell wird, da es Information zueinander in Beziehung bzw. jeweils in einen Kontext setzt, sie also auf eine semantische Ebene hebt und ihnen damit Bedeutung verleiht, auch als „semantisches oder konzeptionelles Schema“ bezeichnet, vgl. (Voß, Jakob, 2009).

Datenmodelle sind meistens unabhängig von einer bestimmten Syntax und lassen sich in verschiedenen Systemen realisieren, sie bilden aber die Verbindung zwischen den Informationseinheiten (A) und einer dritten, syntaktischen Ebene. Diese dritte Ebene drückt das Datenmodell in der anschließenden konkreten Anwendung

³⁹ Unstrukturierte Daten produzieren meistens Unklarheiten, die das Verstehen bzw. die Verarbeitung durch eine Maschine erschweren.

⁴⁰ So gibt es auch eine Vielzahl von Modell-Bezeichnungen: data model, reference model, entity relationship model, abstract model, application profile, meta model, ontology.

entsprechend einer bestimmten Serialisierung bzw. in einem Datenformat aus, setzt es also in konkrete syntaktische Regeln um. Beispiele sind XML oder andere maschinenlesbare Formate.

Datenmodelle unterstützen die Interoperabilität von Daten unterschiedlicher Formate, können also die Nutzbarkeit von Metadaten über verschiedene Systeme hinweg unterstützen, wenn dasselbe Modell zugrunde liegt oder die Modelle interoperabel sind. Metadatenmappings („Crosswalks“) zur Datenkonvertierung zwischen verschiedenen Metadatenstrukturen beschreibt beispielsweise (Getty Research Institute, 2009). Herausforderungen des Zusammenführens von Metadaten unterschiedlicher Herkunft in der Praxis werden weitgehend in der Missachtung von Modelkonformitäten gesehen, d.h. dass Elemente aus Datenmodellen nicht semantisch korrekt verwendet werden, wie beispielsweise in (Keßler & Rühle, 2008) dargelegt.

3.4 Information extrahieren

Ähnlich wie bei der traditionellen Erschließung besteht auch bei der Erschließung von 3D-Modellen die Herausforderung, dass die Modelle seitens der Autoren nur rudimentär mit Metadaten annotiert werden. Obschon Metadaten wie Autor oder Titel bisweilen in wohldefinierter Form vorliegen und somit leicht extrahiert werden können, besteht das Hauptproblem der Erschließung von 3D-Modellen in der Anreicherung um Klassifikationsmerkmale und metatextartige Beschreibungen, d.h. in der inhaltsbasierten Erschließung. Diese nur implizit in den 3D-Daten enthaltenen Informationen, vgl. 3.4.1, erfordern eine elaborierte Herangehensweise zur Extraktion.

Aufgrund des auch in der Domäne der 3D-Modelle stetig steigenden Datenvolumens, das durch die qualitätsorientierte Erschließungskette etablierter Fachbibliotheken nicht länger bewältigt werden kann, werden in diesem Bereich insbesondere Methoden der inhaltsbasierten Erschließung durch vernetzte Nutzergruppen sowie (semi-)automatisierte maschinelle Verfahren stärker nachgefragt, vgl. Darstellung in 1.2.1.

Nach dem Überblick über traditionelle Erschließung und andere Formen der manuellen Metadaten-Produktion in Abschnitt 3.2 führt dieses Kapitel in die maschinelle Erzeugung von Metadaten für 3D-Modelle ein. Dabei werden die Konzepte und Techniken des Multimedia Information Retrievals, des maschinellen Lernens und der Klassifikation hinsichtlich ihres Nutzens für das (teil-)automatisierte Generieren von Metadaten erörtert. Zur Vertiefung wird die Lektüre von (Havemann, Ullrich, & Fellner, 2012), (Lew, Sebe, Djeraba, & Jain, 2006) und (Mølgaard, 2009) empfohlen.

3.4.1 Wissen in 3D-Modellen?

In Abschnitt 2.3 wurde auf strukturelle Eigenschaften von 3D-Modellen eingegangen. Abbildung 11 verdeutlicht mit einem polygonalen Modell beispielhaft die Herausforderung der semantischen Verarbeitung dieser Daten.

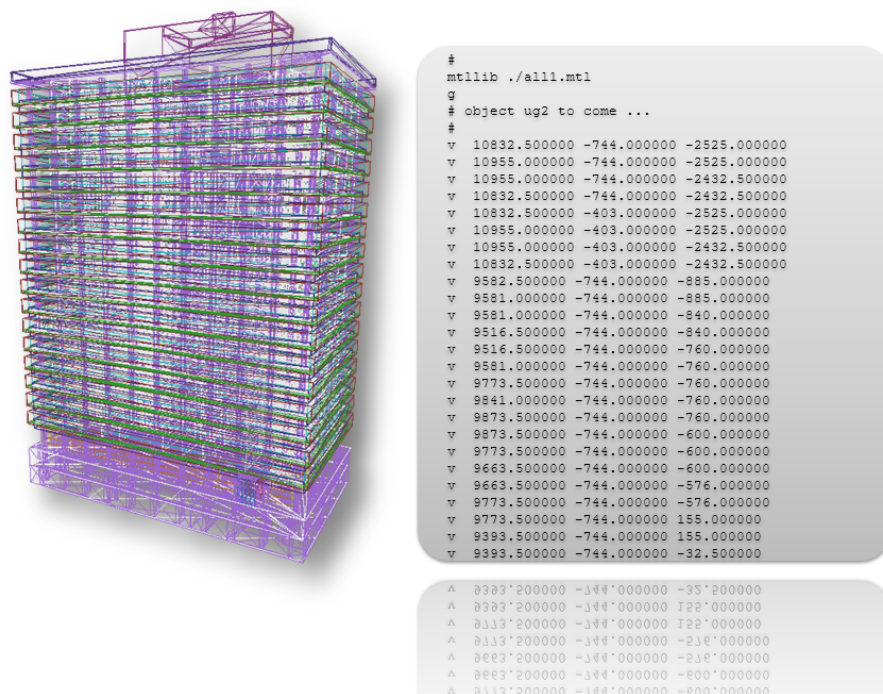


Abbildung 11: Polygonales Modell im Wavefront .obj-Format, geöffnet im CAD- sowie im Texteditor

Auf der linken Seite ist die Darstellung des Drahtmodells eines mehrgeschossigen Gebäudes im CAD-Programm zu erkennen, rechts ein Auszug aus dem Texteditor. Die Einträge im Editor enthalten pro Zeile entweder eine Knotenposition im Raum oder die Definition eines Polygons und stellen in ihrer Gesamtheit das Modell als Gitter dar.

Die Informationen darüber, ob es sich um ein Architekturmodell handelt, um ein Gebäudemodell, um ein Hochhaus oder einen anderen Bautyp, sind zwar vorhanden, jedoch sind die Verarbeitungsmöglichkeiten des menschlichen Gehirns zu eingeschränkt, um aus den gegebenen Zahlen die gewünschte Information zu extrahieren. Dagegen gestattet die visuelle Darstellung im CAD-Programm, gepaart mit dem menschlichen Erfahrungswissen, sofort die Möglichkeit einer Inferenzbildung. Ein Mensch kann aufgrund seines Wissens visuell erfasste Geometrien direkt einem Gegenstand der realen Welt sowie einem oder mehreren Konzepten zuordnen, vgl. Original vs. Modell in 2.1.1. Ein Computer muss die lediglich implizit vorhandene Bedeutung eines 3D-Modells erst lernen, also die Erkennung des durch die Koordinaten repräsentierten Inhalts und die Zuordnung zu Konzepten repräsentierenden Entitäten.

3.4.2 Merkmalsextraktion und Erschließung

Um ein inhaltsbasiertes Information Retrieval anzubieten, sind 3D-Modelldaten hinsichtlich inhaltstragender Merkmale – auch Indexdaten, Features oder Deskriptoren genannt – zu analysieren. Erst die Indexdaten werden im Information Retrieval durch Algorithmen miteinander verglichen, sodass Ähnlichkeiten zwischen ihnen festgestellt werden können. Es ist daher wichtig, konzeptionell zwischen den Rohdaten, den Indexdaten sowie den Metadaten zu unterscheiden⁴¹. Die Me-

⁴¹ Im Gegensatz zu der in der vorliegenden Untersuchung verwendeten Sichtweise der Informatik, die sich aus den dargelegten Prozessierungsschritten ergebenden Trennung der Begriffe Indexdaten und Metadaten ableitet, wird in der informationswissenschaftlichen Literatur das Indexieren je nach Darstellung auch als Zuordnung von Termini zu Informationsobjekten aufgefasst, also synonym zur vorgenannten traditionellen Inhaltserschließung verwendet, vgl. (Umlauf, 2000, Kapitel 1.6.3.3).

tadaten werden erst mit maschinellen Lernverfahren aus den Indexdaten abgeleitet.

In diesem Abschnitt werden die Merkmalsanalyse und die Schritte der 3D-Modellprozessierung von den Rohdaten bis zu den Indexdaten erörtert. Der Prozess der Merkmalsextraktion erfolgt normalerweise in mehreren Stufen. So beschreiben (Bustos Cárdenas, Keim, Saupe, Schreck, & Tatu, 2008) eine Abfolge von der Vorverarbeitung, über die Objektabstraktion und die numerische Umwandlung bis zur Indexdatenerzeugung. Im Folgenden wird die Prozessierung entlang von 7 Schritten erläutert, die sich unter die Stufen Vorverarbeitung, Merkmalsextraktion und inhaltsbasierte Erschließung gruppieren. Weiterhin werden Metadatentypen eingeführt, die an 3 unterschiedlichen Stellen erzeugt werden.

Vorverarbeitung

Abbildung 12 zeigt die ersten Schritte der Modellprozessierung.

Als Eingabe in die Prozessierungskette dienen 3D-Modell-Rohdaten unterschiedlicher Formate. Aus den Rohdaten kann die Extraktion technischer Attribute erfolgen, beispielsweise die Menge an Polygonen, das Dateiformat oder die Dateigröße.

Da die Dateiformate der zu verarbeitenden Modelle meistens sehr heterogen sind, die Indexdaten erzeugenden Algorithmen aber auf jedes Format angepasst werden müssen, werden die Daten in ein einheitliches Format *konvertiert*. Im gleichen Arbeitsschritt können Vorschauformate wie eine 3D-PDF-Vorschau erzeugt werden.

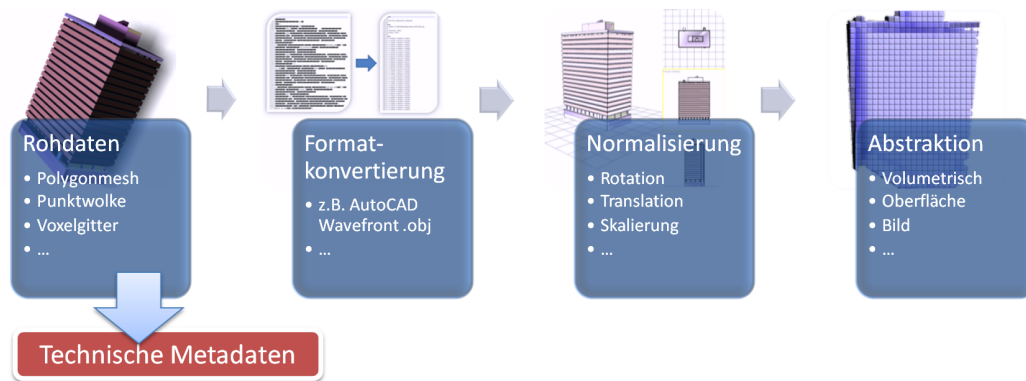


Abbildung 12: Modellprozessierung (Teil 1) - Vorverarbeitungsschritte, Normalisierung, Objektabstraktion

Die sogenannte *Normalisierung*, dient dazu, Invarianz- und Robustheitseigenschaften zu erzeugen, sodass eine Skalierung mit einem bestimmten Maßstab oder eine einheitliche Orientierung möglich ist, z.B. damit die Höhe im Modell immer der Z-Achse entspricht, vgl. Abbildung 12. Die Vorverarbeitung von 3D-Modellen wird modelltypabhängig angewendet und ist nicht immer rein maschinell möglich, vgl. (Berndt, Blümel, & Wessel, 2010, S. 86).

Es folgt die *Objektabstraktion*, d.h. die Überführung des Modells in die gewünschte geometrische Repräsentation, vgl. 2.3.2.

(I. Schmitt, 2005, S. 79) führt bereits nach der Normalisierung die *Segmentierung* an, d.h. eine Teilung des Modells in definierte Einzelabschnitte. Dieser Schritt wird in der Praxis aber eher direkt vor der Erzeugung der Indexdaten angewendet, wie in dem 2. Teil der Prozessierungskette in Abbildung 13 dargestellt.

Merkmalsextraktion

Im nächsten Schritt werden Indexdaten erzeugt. Es gibt folgende Repräsentationen von Indexdaten, siehe auch (Tangelder & Veltkamp, 2008):

- Globale sowie lokale geometrische Eigenschaften, die als (gegebenenfalls im Raum lokalisierte) Vektoren repräsentiert werden; sie beschreiben die *Form*.
- Statistische Deskriptoren, mit denen 3D-Modelle auf spezifische Eigenschaften überprüft werden können; bei 3D-Modellen sind dies z.B. Informationen

über *Farbverteilung* auf der Modelloberfläche, die normalerweise in Form eines Histogramms zusammengefasst werden.

- *Strukturelle* Beschreibungen, die in Form eines Graphen repräsentiert werden.

Aus den Indexdaten lassen sich inhaltlich-technische und quantifizierbaren Attribute erzeugen, z.B. die Größe der Raumflächen.

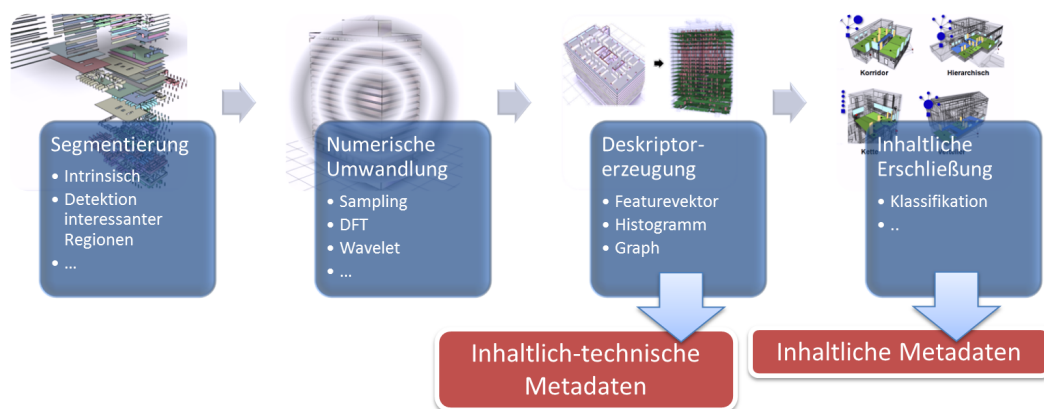


Abbildung 13: Modellprozessierung (Teil 2) - Segmentierung, Umwandlung sowie anschließende Merkmalsextraktion und Erschließung

Inhaltsbasierte Erschließung

Bei inhaltsbasierten maschinenunterstützten Erschließungsmethoden wird auf maschinelle Lernverfahren zurückgegriffen, deren Umsetzung meist durch Lernalgorithmen erfolgt. Maschinelles Lernen ist eine aus der Künstlichen Intelligenz entlehnte Methode und ermöglicht die „künstliche“ Generierung von Wissen aus „Erfahrung“, die der Maschine antrainiert werden kann. Somit ist sie ähnlich wie der Mensch in der Lage, Inferenzen zu bilden, und es wird, beispielsweise aufgrund von Form- oder Strukturmerkmalen, eine automatische Kategorisierung von 3D-Modellen anhand gegebener Schemata ermöglicht.

Die Klassifikation von 3D-Modellen, also die automatische Einteilung von Modellen in eine vordefinierte Menge von Klassen, wird durch überwachtes Lernen⁴² ermöglicht. Der Maschine wird dazu ein Satz von Beispielen (Trainingsdaten) gegeben, wobei jedes Beispiel aus einem Modell und dem zugehörigen Klassenlabel besteht. Der erste Schritt vor der eigentlichen Kategorisierung, die darin besteht, unter Zuhilfenahme dieser Beispiele eigenständig Vorhersagen für neue Beispiele (Testdaten) zu generieren, ist die Extraktion von Merkmalen, die die einzelnen Klassen voneinander unterscheiden (siehe 3.4.2). Dabei versuchen Diskriminanzfunktionen anhand dieser Merkmale zu klassifizieren. Ziel ist, gute Features zu bauen, also solche, die diskriminieren.

Benchmarks sind Testdatenbanken mit manuell vorklassifizierten Objekten, beispielsweise 3D-Modellen, die sich durch die Anwendungsdomäne, die geometrischen Repräsentationen der 3D-Modelle und den Typ des Klassifikationsschemas, das für den Benchmark entwickelt wurde, unterscheiden. Weitere Merkmale, über die vor allem auf die Anwendbarkeit eines Benchmarks für das maschinelle Lernen geschlossen werden kann, sind die Granularität des Klassifikationsschemas und die Gleichmäßigkeit der Verteilung von Modellen auf die einzelnen Klassen. Benchmarking ermöglicht es, die relative Leistungsfähigkeit von Retrieval- und Klassifikationsalgorithmen untereinander objektiv zu vergleichen. Zusätzlich kann evaluiert werden, wie gut die Algorithmen hinsichtlich der durch die entsprechenden Klassifikationsschemata definierten Anforderungen in der spezifischen Domäne performieren. Mit überwachten Lernmethoden kann eine Verbesserung der Ergebnisse für eine Suchanfrage sowie eine automatische Klassifikation von Objekten erzielt werden. In den letzten Jahren wurden zahlreiche Benchmarks für 3D-Modelle geschaffen. Relativ bekannt ist der Princeton Shape Benchmark, der 3D-Modelle unterschiedlicher Themengebiete (z.B. Tiere, Pflanzen, Objekte) ent-

⁴² Maschinelles Lernen kann überwacht oder unüberwacht sein. Beim unüberwachten Lernen muss die Maschine selbstständig eine Struktur aus den Daten extrahieren, also eine Menge von Eingabeobjekten in Gruppen clustern, sodass Objekte innerhalb einer Gruppe ähnlich zueinander sind (z.B. hinsichtlich ihrer Geometrie). Der wesentliche Unterschied zum überwachten Lernen ist, dass der Maschine keine vorklassifizierten Beispieldaten zur Verfügung stehen.

hält, vgl. (Shilane, Min, Kazhdan, & Funkhouser, 2004). Innerhalb des Projekts PROBADO, vgl. (Berndt, Blümel, Clausen, u. a., 2010), wurde ein Benchmark exklusiv für Objektmodelle aus dem Fachbereich Architektur entwickelt (Wessel, Blümel, & Klein, 2009). Beispiele für andere Medienformate finden sich beispielsweise in (Petras & Clough, 2011): CLEF.

3.4.3 Typen automatisch extrahierter Metadaten

Im Vorstehenden wurden anhand von traditionellen und anhand von maschinellen Erschließungsmethoden die Prozesse zur Erzeugung von Metadaten erörtert. Metadaten können somit als formale und inhaltliche Aussagen über Informationsobjekte bezeichnet werden, die den Objekten manuell zugeordnet oder ihnen automatisiert entnommen wurden.

Automatisch können Metadaten an insgesamt drei Stellen erzeugt werden, wie in Abbildung 12 und Abbildung 13 illustriert. An der dargestellten Prozesskette werden unterschiedliche Metadatatypen verortet.

- Dateiattribute, also technische Metadaten, fortan bezeichnet als „low-level“ Metadaten,
- technisch-inhaltliche Metadaten, die meistens quantifizierbare 3D-Modellattribute beschreiben, fortan bezeichnet als „medium-level“ Metadaten,
- inhaltliche Metadaten, die mittels automatischer Klassifikation erzeugt werden, fortan bezeichnet als „high-level“ Metadaten.

Bei der Entwicklung des Metadatenmodells neben den in Abschnitt 2.5 genannten Anforderungen ist weiterhin zu beachten, dass im Hinblick auf die Entwicklung neuer Extraktionsmethoden und Algorithmen eine Erweiterbarkeit um Attribute aller drei Metadatatypen möglich sein muss.

3.5 Formalisierte Annotation im Linked Data

Nach der Definition der Eigenschaften von Metadaten soll ein Exkurs zu den Ideen und Methoden des Semantic Web sowie der aktuellen Ausprägung, dem Linked Data, Aufschluss darüber geben, welche Anforderungen und Möglichkei-

ten aktuell zur Repräsentation und Kontextualisierung digitaler Ressourcen vorliegen.

3.5.1 Web of Documents → Web of Data

Das WWW, auch als „Web of Documents“ bezeichnet, verlinkt Informationen auf Dokumentenebene miteinander. Der Benutzer hat es, wie als Herausforderung in Kapitel 1 angesprochen, mit einer großen, unsortierten Flut von Informationen innerhalb unterschiedlicher Medienformate wie Texten, 2D- oder 3D-Inhalten, Audio, Video, etc. zu tun. Das Herauspicken der relevanten Information gestaltet sich als langwieriger und aufwändiger Prozess. Die unstrukturierte bzw. semi-strukturierte Form (z.B. in HTML) sowie das Fehlen einer maschinell prozessierbaren Bedeutungsdefinition zur jeweiligen Informationen erschweren Computern das Finden relevanter Information zur Suchanfrage, da sie im Gegensatz zum Menschen kein Kontextwissen haben.

Das in 2001 erstmals erwähnte Semantic Web wird als eine Weiterentwicklung des WWW (als Verknüpfung von Dokumenten) gesehen. (Berners-Lee, Hendler, & Lassila, 2001) definieren

“The Semantic Web is not a separate Web, but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation.”

Die Definition betont, dass Informationen im WWW mit einer genau definierten Bedeutung versehen werden. So setzt das Semantic Web strukturierte Annotationen bzw. Metadaten voraus. Die Perspektive ist, sämtliche Information strukturiert, d.h. maschinenlesbar im WWW verfügbar zu haben und die Informationen in den Dokumenten mit einer eindeutigen Beschreibung ihrer Bedeutung versehen, die auch von Computern verarbeitet werden kann und, darüber hinaus, durch Inferenzbildung auch „verstanden“⁴³ werden soll. Ein Umdenken in Richtung Linked Open Data setzte vor ca. 6 Jahren ein: Ziel ist es seitdem, die bereits vorhandenen strukturierten Daten ins Netz zu bringen und interoperabel zu machen,

⁴³ zu „Verstehen“ siehe 3.2.1. Siehe außerdem Ausführungen zur Hermeneutik in (Capurro, 2006).

anstatt dass man versucht – der originären Idee des Semantic Web folgend – die schon im Netz befindlichen Daten mit einer Bedeutung zu versehen.

Berners-Lee, der Erfinder des WWW, gab bei einem Vortrag an der Princeton Universität in den USA im Jahre 2006 zu, dass es vielleicht falsch war, den Namen „Semantic Web“ für seine Idee zu wählen. Stattdessen wird die Idee als „Web of Data“, da es sich um strukturierte, verlinkte Daten handelt, oder als „Web of Concepts“ umschrieben (Berners-Lee, 2006a, S. 21).

3.5.2 Umsetzung

Kurzfassen könnte man die Linked Data-Definition als Bereitstellung und Nutzung von Daten über http-URIs in einem RDF⁴⁴-Format. Die von Tim Berners-Lee aufgestellten Linked-Data-Prinzipien⁴⁵ dienen als anerkannte Konventionen für die Publikation von Linked Data:

- “1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL)
4. Include links to other URIs, so that they can discover more things.”

Es wurde bereits eine große Anzahl Daten entsprechend dieser Linked-Data-Prinzipien publiziert. Aktuelle Darstellungen der gemäß der Linked-Data-Prinzipien veröffentlichten Datensammlungen werden ständig aktualisiert und sind verlinkt auf der Seite des W3C / LOD⁴⁶ zu finden. Da bereits unterschiedlichste Literatur zu Linked Data und seiner Umsetzung in div. Anwendungskontexten existiert, vgl. beispielsweise (Heath & Bizer, 2011) oder (Hyvönen, 2012),

⁴⁴ Das Resource Description Framework (RDF) dient als Modell zum Datenaustausch im WWW, vgl. <http://www.w3.org/RDF/>.

⁴⁵ <http://www.w3.org/DesignIssues/LinkedData.html>

⁴⁶ <http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData>

wird an dieser Stelle nur kurz auf die für eine Umsetzung notwendigen Facetten eingegangen:

Wichtige Idee im Kontext Linked Data ist der Schwenk von einer Tabellensicht (Ausgang sind Daten in den genannten „Datensilos“) auf eine Aussagensicht. Die Bestandteile einer Aussage – Subjekt, Prädikat und Objekt – müssen mittels HTTP-URIs eindeutig bezeichnet sein, wobei das Objekt auch als Literal dargestellt werden kann. Durch die Dereferenzierbarkeit der URIs erhält man die Beschreibung der Ressource bzw. des Vokabulars, aus dem das Prädikat stammt. Bei den durch Präfixe abgekürzten URIs spricht man auch von den Namensräumen (Namespaces), in denen Ressourcen und Prädikate existieren. Die Klassen und Eigenschaften, die verwendet werden, können in Beschreibungssprachen für Vokabulare definiert werden, z.B. RDF Schema (RDFS), oder Web Ontology Language (OWL). Weiterhin wird festgelegt, welche Beziehungen zwischen Instanzen dieser Klassen herrschen können. So ist es möglich, durch die auf Vokabularebene gegebene Information, implizite Aussagen auf Daten-Ebene zu inferieren.

RDF-Schemata sind gut dafür geeignet, Daten aus verschiedenen Quellen zu integrieren, wenn unterschiedliche Vokabulare verwendet werden.⁴⁷

RDF dient damit prinzipiell als Metamodell, um Interoperabilität zu erreichen und (Bibliotheksum-) daten auf das Wesentliche zu reduzieren. Im Hinblick auf die in Abschnitt 1.2.2 dargestellte Verschiebung des Bibliotheksumfeldes durch das WWW bewerten (Dunsire, Hillmann, Phipps, & Coyle, 2011) das Thema Mapping-Strategien und Interoperabilität von Metadaten-Modellen, vgl. 3.3, neu:

“The meaning of “mapping” changes radically on moving from a database and record based approach to an open, multi-domain, global, shared environment based on linked data technologies -- where [...] a nearly infi-

⁴⁷ Die (ob der mittlerweile umfassenden Literatur sowie Bekanntheit des Themas Linked Data in der Bibliothekswelt möglichst kompakt gehaltenen) Inhalte dieses Absatzes sind an (Ostrowski & Christoph, 2011) angelehnt.

nite number of properties can be defined to describe an infinite number of entities, and authority is multi-dimensional and often ephemeral.“

und schlagen eine Änderung des Vorgehens in der Repräsentation vor: “from record-based to statement-based, and from syntax-based to semantic-based“. Die referenzierten Linked-Data-Technologien werden als Basis für ein grundlegend anderes, tendenziell flexibles und offeneres Datenmodell beschrieben.

3.5.3 Ausgewählte Modelle und Vokabulare

Bevor im nächsten Kapitel ein Metadatenmodell für 3D-Modelle aus den definierten Anforderungen abgeleitet und spezifiziert wird, sollen abschließend Metadatenmodelle mit Potential für eine kontextualisierte Repräsentation von 3D-Modellen erörtert werden. Der Schwerpunkt liegt dabei bei den generischen Modellen, weil diese das Potential haben, als „Dächer“, also übergeordnete Modelle, über den Modellen der einzelnen Communities zu stehen.

DC / DCterms und weitere gebräuchliche Vokabulare

Das bekannteste Beispiel für deskriptive Metadaten ist das einfache Dublin Core Metadata Element Set (DC) der Dublin Core Metadata Initiative (DCMI)⁴⁸. Es wurde 1995 zunächst zur Beschreibung von WWW-Ressourcen entwickelt und ist ISO-zertifiziert. Unter DC ist zunächst ein Minimalsatz von 15 „Basis“-Metadatenelementen für „dokument-ähnliche Objekte“ aufzuführen, die die Vorstellung von Metadaten bis heute prägen. Sie stellen einen domänenübergreifenden Konsens für Metadaten im Sinne von in Katalogen gebräuchlichen „Record Sets“ dar. Demnach handelt es sich bei Dublin Core weniger um den Versuch, ein Datenmodell (der realen Welt, s.o.) einzuführen als vielmehr um eine „flache“ Metadatenstruktur. Die Dublin Core -Metadatenfelder sind optional, können mehrfach und in beliebiger Reihenfolge auftreten. Das Set wird zum einfachen Austausch von Metadaten als elementare Beschreibungen von Informationsobjek-

⁴⁸ Das DCMI ist eine Organisation zur Verbreitung interoperabler Metadaten-Standards und die Entwicklung spezieller Metadaten-Vokabulare zur Beschreibung von Ressourcen, vgl. <http://dublincore.org/>.

ten angewendet und meist mit HTML-Tags oder unter Verwendung von XML-Schemata encodiert.

Während das „einfache“ DC die Elemente des Metadatensets nur als Element-Value-Paare benutzt, stellen die DCMI Metadata Terms (DCterms) zusätzliche sowie detaillierende Felder (element refinements) bereit, die eine speziellere Beschreibung bzw. Kategorisierung erlauben. Die Definition von Domänen und Ranges erlaubt die Einbeziehung anderer Semantic Web Vokabulare und den Gebrauch im Linked Data Umfeld.

Das Simple Knowledge Organisation System (SKOS) dient zur öffentlichen Bereitstellung und Kombination kontrollierter, strukturierter und maschinenlesbarer Vokabulare als Linked Data, siehe beispielsweise (Eckert, 2011).

CIDOC CRM und FRBR_{oo}

CIDOC CRM ist ein objektorientiertes, ereigniszentriertes und erweiterbares Modell zur Beschreibung von Begriffen und Informationen im Bereich des Kulturerbes. CIDOC unterstützt RDF⁴⁹

FRBR ist ein konzeptionelles Entity-Relationship Modell, das von der International Federation of Library Associations and Institutions (IFLA) 1998 publiziert wurde, vgl. (IFLA Study Group on the Functional Requirements for Bibliographic Records & International Federation of Library Associations and Institutions, 1998). Daraufhin gab es zwar theoretische Diskussionen in der Fachwelt; konkrete Umsetzungen in Software und Formaten blieben weitgehend aus.

Mit der Einführung des Standards Resource Description and Access (RDA) in Europa gerät FRBR wieder in den Fokus. Die FRBR_{oo}⁵⁰-Initiative aus den Arbeitsgruppen und CIDOC CRM und FRBR hat zum Ziel, FRBR objektorientiert und auf Basis der CRM-Modellierung zugrunde liegenden Konzepte zu modellieren, vgl. (Bekiari, Doerr, & Bœuf, 2012, Kapitel 1.1.2). Die beiden objektorien-

⁴⁹ http://www.cidoc-crm.org/rdfs/cidoc_crm_v5.0.4_official_release.rdfs

⁵⁰ object orientented FRBR

tierten Modelle FRBROO und CRM ergänzen sich insbesondere durch die Vereinigung von Klassen und Properties aus den Kontexten Bibliothek und Kulturerbe.

3.5.4 Stand der Forschung / Verwandte Arbeiten: Projekte und ihre Modelle

Im Projekt 3D-COFORM⁵¹ wurde versucht, Lösungen für Metadaten von 3D-Modellen des Kulturerbes unter Verwendung von Linked Data Technologien zu entwickeln, vgl. (Felicetti, D'Andrea, & Niccolucci, 2012).

Das FAÇADE (Future Proofing Architectural Computer Aided-Design)-Projekt⁵² (2006-2009) entwickelte ein Project Information Model mit einer Metadata Ontologie für proprietäre Standards (incl. Revit etc., allerdings kein IFC), vgl. (Smith, 2009).

Das MACE (Metadata for Architectural Content in Europe)-Projekt⁵³ entwickelte klassifikatorische Organisationssysteme basierend auf AAT und an LOM angelehnte Metadaten.

Die Europeana⁵⁴ besitzt als Virtuelle Bibliothek für das wissenschaftliche und kulturelle Erbe Europas das Europeana Data Model (EDM). Es nutzt RDF, und basiert im Wesentlichen auf OAI-ORE, SKOS und Dublin Core, vgl. (Europeana v1.0 Group, 2012).

Das CARARE-Projekt (-2012) beschäftigte sich als eines der Europeana-Satellitenprojekte⁵⁵ mit der Erschließung und Integration von digitalen Ressourcen aus dem archäologischen und architektonischen Kulturerbe, speziell 3D-Modellen.

⁵¹ <http://www.3d-coform.eu/>

⁵² <http://facade.mit.edu/>

⁵³ <http://portal.mace-project.eu/>

⁵⁴ <http://www.europeana.eu/portal/>

⁵⁵ <http://carare.eu/News/CARARE-and-EDM-Case-Study>

Lösungen für Metadaten aus dem Bereich Langzeitarchivierung von Objekten des Kulturerbes entwickeln (Beaudoin, 2012a) bzw. (Beaudoin, 2012b).

4 Metadaten für 3D-Modelle

4.1 Vorbemerkung

Die Erstellung des Metadatenmodells für 3D-Modelle stützt sich auf die in Kapitel 2 und 3 erläuterten Anforderungen. Das Metadatenmodell dient als Integrationsmedium, um 3D-relevante Information abzubilden, zu verbinden und somit die 3D-Modelle anzureichern. Es werden inhaltlich relevante und unterscheidbare Informationseinheiten definiert, in deren Kontext das einzelne 3D-Modell zu setzen ist.

Diese Herangehensweise hat im Gegensatz zu der Auffassung von Modelldatei-Metadaten als „Records“ das Potenzial, Beziehungen zu weiteren Modellen, anderen Medienformaten bzw. Dokumenten und modellrelevanten Informationseinheiten bzw. Entitäten herzustellen, welche den Informationskontext zum Beschreiben und letztendlich zum Erleichtern des Retrievals des Modells herstellen. In Anbetracht selten umfassend vorliegender Metadaten ist die Einführung komplizierter, „endloser“ Modellkontexte weder nachvollziehbar noch notwendig. Vielmehr wird vom Modellkontext nur das dargestellt, was unbedingt für die Kontextualisierung des 3D-Modells selbst relevant ist und dessen Auffinden dient.

Ausgegangen wird hierbei von den in Abschnitt 2.4.2 dargelegten Fallbeispielen bzw. Suchstrategien. So kann beispielsweise ein Benutzer mit Kenntnis des Architekten eines real existierenden Gebäudes, welches in 3D nachmodelliert wurde, durch die Entität CONTRIBUTOR (vgl. 4.5.1), verbunden mit dem realen Objekt (vgl. 4.6.2) direkt zum 3D-Modell gelangen. Weiterreichende Informationen zum Gebäude, wie beispielsweise der Entstehungskontext desselben, können detailliert in den durch entsprechend angebrachte URIs verlinkten, spezifischen Wissensbasen respektive Ontologien abgebildet werden. An den entsprechenden Stellen werden daher Hinweise auf domänenspezifisch aktuell sinnvoll zu verlinkende Begriffssystematiken gegeben, die zur Kontextualisierung und Einbettung in das Web of Data dienen sollen, vgl. Abschnitt 4.3.2.

Zusammenfassend kann gesagt werden, dass das 3D-Modell selbst naturgemäß das Zentrum der Beschreibung bildet und eine möglichst umfassende Kontextuali-

sierung ermöglicht wird. Gleichzeitig wird mit wenigen Entitäten, flexiblen Attribut-Typen sowie minimalen Annotationsanforderungen pro Modell der bereits diskutierten höchst heterogenen Menge an verfügbaren Metadaten Rechnung getragen und eine möglichst flexible Erweiterbarkeit für Informationsextraktion inklusive elaborierter Methoden der inhaltsbasierten Erschließung wie beispielsweise Autoklassifikation gewährleistet.

4.2 Namensräume

Das Modell wurde zunächst als Entity-Relationship-Modell (ER-Modell) für die Umsetzung in einer relationalen Datenbank entwickelt. Daher folgt die Darstellung im ersten Teil dieses Kapitels diesem Ansatz. Hauptentitäten können prinzipiell wie Klassen einer objektorientierten Modellierung betrachtet werden, die Attribute und Relationen wie Properties. Die Einführung der Entitäten erfolgt unter Berücksichtigung der jeweiligen Vorgaben für notwendige Entitäten und Attribute aus den vorigen Kapiteln, vgl. 4.1, auf die an den entsprechenden Stellen referenziert wird. Beispiele sollen die inhaltliche Nachvollziehbarkeit erleichtern.

Außerdem werden Äquivalente zu Elementen anderer Modelle und Ontologien gesucht. Von den beiden Möglichkeiten, sich entweder bestehender Ontologien für das Namespace-Alignment zu bedienen, oder sämtliche Klassen und Properties neu zu erfinden, wird der erstgenannten aus Interoperabilitätsgründen der Vorzug gegeben. Es wird also dem eher pragmatischen Ansatz gefolgt, möglichst viele gebräuchliche bzw. etablierte Namespaces (optimalerweise W3C-standardisiert) für das eigene Modell zu adaptieren, wobei auf konsistentes Alignment geachtet wird. Das Ziel ist eine *möglichst komplette* Abbildung der 3D-domänspezifisch relevanten Aussagen durch Aufbau auf *bereits vorhandene* generische Modelle, die explizit zur Erweiterung ausgelegt sind. Insbesondere das momentan in der Entwicklung befindliche FRBRoo stellt Elemente bereit, die zur vollen Ausdrucksstärke eines Großteils der 3D-Modell-beschreibenden Informationsfacetten verwendet werden können. Ziel von FRBRoo ist es, FRBR objektorientiert und auf Basis der der CRM-Modellierung zugrunde liegenden Konzepte zu modellie-

ren.⁵⁶ Da für FRBRoo zum Abgabezeitpunkt dieser Arbeit noch keine Namensraum-Adresse⁵⁷ vorliegt, werden die aktuell unter (Bekiari u.a., 2012) verzeichneten Klassen und Property-Bezeichner übernommen (F: FRBRoo-Klassen-, R: FRBR-Properties-, E: CRM-Klassen-, P: CRM-Properties-Deklaration).

EDM, das ähnlich wie das hier eingeführte Modell zur Modellierung viele Namenräume bestehender Ontologien wie SKOS benutzt, stellt z.T. ebenso gut anwendbare Klassen und Properties wie FRBRoo bereit. Das in dieser Arbeit entwickelte Modell konzentriert sich aber aufgrund des o.g. möglichst konsistenten Alignments auf FRBRoo, zumal momentan innerhalb einer „EDM – FRBRoo Application Profile Task Force (EFAP-TF)“⁵⁸, ein Abgleich von EDM zu FRBRoo geschaffen wird.

Über FRBRoo hinaus nutzt das Modell (hauptsächlich auf Seiten der Prädikate) die folgenden Namensräume nach:

- das Resource Description Framework (RDF) und den RDF Schema (RDFS) Namespace,⁵⁹
- den Simple Knowledge Organization System (SKOS) Namespace,⁶⁰
- den Dublin Core Namespace for Elements (abgekürzt als dc)⁶¹, Terme, abgekürzt als dcterms),⁶²
- den Namespace des Friend of a Friend Vokabulars (FOAF).⁶³

⁵⁶ "Aligning (possibly even merging) the two object-oriented models", d.h. FRBRoo und CRM ergänzen sich.

⁵⁷ Für die CRM-Klassen und Properties liegt bis dato folgende Adresse vor: http://www.cidoc-crm.org/rdfs/cidoc_crm_v5.0.4_official_release.rdfs.

⁵⁸ <http://pro.europeana.eu/web/network/europeana-tech/-/wiki/Main/Task+Force+EDM+FRBRoo>

⁵⁹ <http://www.w3.org/TR/rdf-concepts/>

⁶⁰ <http://www.w3.org/TR/skos-reference/>, <http://www.w3.org/2009/08/skos-reference/skos.html>

⁶¹ <http://purl.org/dc/elements/1.1/>

⁶² <http://purl.org/dc/terms/>

⁶³ <http://xmlns.com/foaf/0.1/>

Bei einigen 3D-spezifischen Attributen, die durch die o.g. Namensräume nicht abgedeckt werden konnten, werden mittels Falcons⁶⁴ ermittelte Spezialontologien vorgeschlagen. Es kann festgestellt werden, dass bisher ein Quality-Framework für die nach den Designprinzipien des Linked Data bereitgestellten Daten bzw. speziell für Knowledge Organisation Systems (KOS) fehlt, möglicherweise deshalb, weil nach wie vor möglichst viele Datensets publiziert (und vernetzt) werden sollen.

Die Namespaces / Äquivalente dienen außerdem für eine graphenbasierte Modellierung der Entitäten einschließlich ihrer Attribute und Relationen sowie für die Visualisierung anhand eines beispielhaften Gebäudemodells im zweiten Teil des Kapitels. Hier wird insbesondere visualisiert, wie die unterschiedlichen Typen der mittels automatisierter, z.T. inhaltsbasierter Verfahren ermittelten Metadaten, vgl. 3.4.3, integriert werden.

Im Anhang findet sich eine detaillierte, der aktuellen Umsetzung des Modells in der relationalen Datenbank folgende tabellarische Darstellung inkl. Datentypen. Eine prototypische Umwandlung der relationalen Darstellung in RDF wurde bereits mittels D2R⁶⁵ gelöst, aber aufgrund des Umfangs und in Anbetracht der Redundanz zur inhaltlichen Modelldarstellung innerhalb dieses Kapitels nicht aufgenommen. Stattdessen ist zur Veranschaulichung weiterer Inhalte, die durch das 3D-Metadatenmodell abgebildet werden können, aber aus Übersichtsgründen nicht in die graphenbasierte Darstellung aufgenommen werden konnten, eine in XML codierte Beispielausgabe aus der bestehenden OAI-PMH-Schnittstelle im Anhang aufgeführt.

⁶⁴ <http://ws.nju.edu.cn/falcons/ontologysearch/index.jsp>

⁶⁵ <http://d2rq.org/d2r-server>

4.3 Entitäten

4.3.1 Gruppen

Die wichtigsten Konzepte für die kontextualisierte Darstellung von 3D-Modellen wurden bereits in Kapitel 1 identifiziert und gruppiert. Abbildung 14 zeigt schematisch die wichtigsten Entitäten des Modells mit ihren Verbindungen untereinander sowie ihrer jeweiligen Zugehörigkeit zu einer Gruppe.

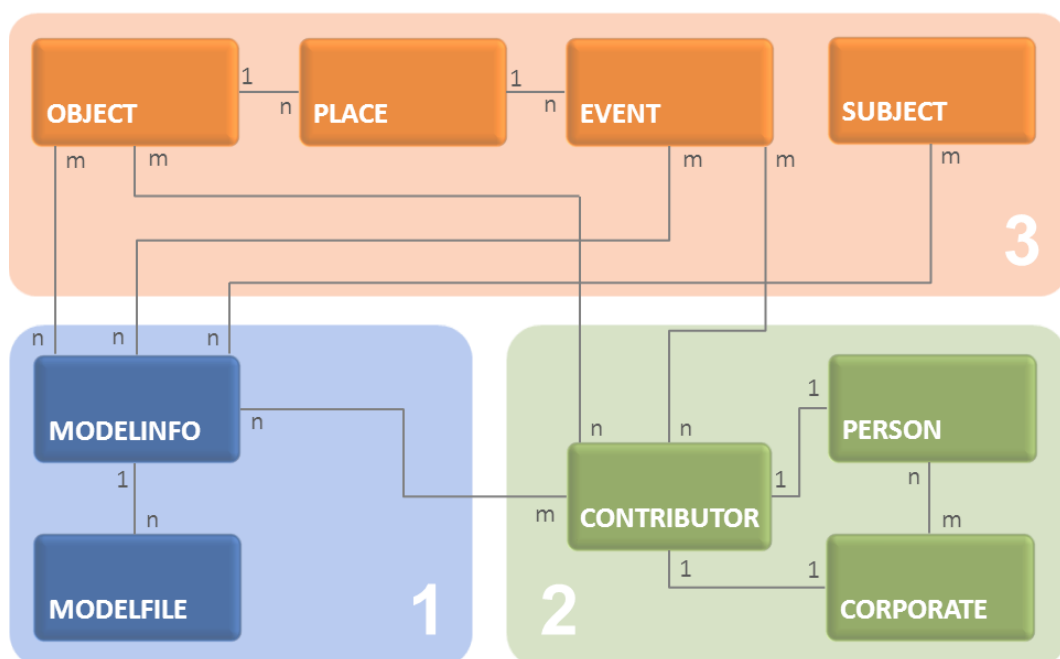


Abbildung 14: Schematische Darstellung der wichtigsten Entitäten und Relationen des 3D Metadatenmodells

- Gruppe 1 beschreibt die 3D-Modelle selbst.
- Gruppe 2 beinhaltet die involvierten Personen oder Körperschaften.
- Gruppe 3 beinhaltet reale Gegenstände / Gebäude, Orte, mit der Modellerstellung in Zusammenhang stehende Ereignisse sowie Begriffe / Konzepte.

Als ursprüngliches Referenzmodell bei der Entwicklung diente FRBR, das erst durch FRBRoo in die Welt der semantischen Integration mit anderen Domänen – wie architektonischen 3D-Modellen – gehoben wird. Der vierstufige Ansatz der

ursprünglichen FRBR-Gruppe 1 wird entsprechend der unter 4.4.1 genannten aktuell stattfindenden Diskussion eines Herunterbrechens der vier Levels auf zwei nicht übernommen. Im Gegensatz zu FRBR werden die Entitäten der Gruppe 3 nicht nur als Themen eines Werkes behandelt.

4.3.2 Einbindung externer Informationsquellen

Eine elementare Entität (im Modell bezeichnet durch EXTERNALINFO) lässt sich keiner Gruppe explizit zuordnen, da sie mit Entitäten aus unterschiedlichen Gruppen in Verbindung stehen kann. EXTERNALINFO kann eindeutig identifizier- und verknüpfbare Information außerhalb der 3D-Metadaten aufnehmen, die analog der Linked Data Designprinzipien von externen Bereitstellern zur Verfügung gestellt wurden und via URIs abgerufen werden können. Hierbei sind neben der DBPedia⁶⁶ insbesondere Normdaten, interessant, welche bereits als Linked Open Data vorliegen:

- ein Personennormsatz (VIAF⁶⁷, GND⁶⁸, ..),
- ein Schlagwortnormsatz (DDC⁶⁹, GND, LCSH⁷⁰, Getty AAT⁷¹, ..),
- eine Ortsentität (Getty TGN⁷², GeoNames Ontology⁷³, ..).

Diese können zur erweiterten Beschreibung und Identifikation herangezogen werden, beispielsweise bei:

- realen und imaginären Objekten (z.B. Bauwerken, Designklassikern, kulturellen Artefakten, ..),

⁶⁶ <http://dbpedia.org>

⁶⁷ <http://viaf.org/>

⁶⁸ http://www.dnb.de/DE/Standardisierung/Normdaten/GND/gnd_node.html

⁶⁹ <http://deweysearchde.pansoft.de/webdeweysearch/mainClasses.html>

⁷⁰ <http://id.loc.gov/authorities/subjects.html>

⁷¹ <http://www.getty.edu/research/tools/vocabularies/aat/>

⁷² <http://www.getty.edu/research/tools/vocabularies/tgn/>

⁷³ <http://www.geonames.org/ontology/documentation.html>

- den jeweiligen Erstellern derselben (Architekt, Designer, Künstler, ..)
- Orten, an denen diese Objekte stehen,
- Ereignissen, in deren Kontext das 3D-Modell erstellt wurde,
- Orten, an denen Ereignisse stattfinden,
- an Ereignissen beteiligten Personen (z.B. Auslober),
- Erstellern / Rechteinhabern / o.ä. des 3D-Modells.

Eine Verbindung der elementaren Entität `EXTERNALINFO` mit „`edm:NonInformationResource`“ aus dem Europeana Data Model ist prinzipiell möglich und sinnvoll.

4.3.3 Darstellung

Im folgenden Abschnitt werden die Entitäten aufgelistet und detailliert diskutiert. Die Auflistung erfolgt der besseren Übersicht halber gruppenweise mit den jeweils übergeordneten Entitäten zuerst.

Über Attribute und Relationen können den Entitäten Eigenschaften zugeordnet werden, die jeweils unter der Entität aufgelistet sind. Tabellarisch dargestellt werden (Occurrence /) Obligation⁷⁴, Beschreibung, akzeptierte Werte und / oder Beispiele sowie adaptierbar Namespace als äquivalente Klassen bzw. Properties inhaltlich geeigneter und relevanter Datenmodelle sowie Vokabularien. Zur Herleitung der Entität aus den Anforderungen der vorangegangenen Kapitel sind einleitende Absätze eingefügt, teilweise ebenfalls bei den Attributen und Relationen.

Übergeordnete Entitäten, vgl. Abbildung 14, werden – durch Nummerierung unterscheidbar – in jeweils eigenen Abschnitten aufgeführt. Die Entitäten haben quasi einen Klassenstatus, wobei alle darunter aufgeführten Attribute und Relationen wie Properties zu behandeln sind, auch wenn sie im ER-Modell weitere Entitäten beinhalten, welche aufgrund der Limitierung einer relationalen Modellierung eingeführt wurden, z.B. die als Property von `MODELFILE` zu behandelnde Entität `SOFTWARE`. Bis auf `MODELINFO` werden die Entitäten-IDs nicht aufgeführt

⁷⁴ (Häufigkeit /) Notwendigkeit des Auftretens

und finden sich im Anhang. Sämtliche für die Implementation in der Datenbank festzulegende Occurrences werden ebenfalls erst in der relationalen Darstellung im Anhang eingeführt. Die IDs zu Modelinfo werden in der relationalen Darstellung auf Seiten der verknüpften Entitäten oder in Zwischentabellen gelöst.

Entsprechend der definierten Gruppen werden die Entitäten genannt, darunter ihre (in der oo Darstellung wie Properties behandelten) Attribute und Relationen zu anderen Entitäten. Die Relationen werden jeweils auf beiden Seiten genannt.

Minimal notwendige Entitäten für die Beschreibung eines Modells sind MODELINFO, CONTRIBUTOR und MODELFILE.

Die Darstellung der jeweiligen Entität / des Attributs / der Relation erfolgt in folgendem Format:

- Eine einleitende textliche Definition sowie deren aus den vorangegangenen Kapiteln abgeleitete inhaltliche oder technische Notwendigkeit,
- „Namespace“ beinhaltet Äquivalente aus anderen Modellen und Ontologien,
- „Beispiele“ beinhaltet eine Liste an Beispielinstanzen sowie, falls es Einschränkungen gibt, mögliche Wertebereiche (eine umfassende Darstellung jeweils akzeptierter Werte findet sich am entsprechenden Attribut im Anhang) und sonstige Limitationen,
- „Occ.“ beinhaltet Informationen zur Häufigkeit und Obligation der Angabe.

4.4 Gruppe 1 – Modelle und ihre Instanzen

4.4.1 MODELINFO

Die zentrale Entität zur Beschreibung intellektueller 3D-Modelleinheiten, welche als abstrakte Darstellung eines 3D-Modells erörtert werden kann.

Namespace:	F2 Expression ⁷⁵ (frbroo:PublicationExpression)
------------	---

Von hier ausgehend werden bei der Beschreibung einzelner 3D-Modelle und ihrem Kontext Relationen zu weiteren Entitäten gebildet, insbesondere zu MODELFILE, Modelldateien unterschiedlicher Formate, welche aus MODELINFO instanziiert bzw. ausgegeben werden können.

Ein detailliertes Gebäudemodell enthält normalerweise weitere Modelle, die extrahiert und als separates Modell nachgewiesen werden können. Ebenfalls ist die Teilung in Informationseinheiten beispielsweise zur Gebäudebewirtschaftung (Raumbücher, unterscheidbare Wartungseinheiten usf.) z.T. notwendig. Analog dieser Schachtelung ist die Rekursivbeziehung von MODELINFO möglich.

Die folgende Übersicht beschreibt in kompakter Form die zentralen MODELINFO-Attribute sowie assoziierter Entitäten. Die Reihenfolge der Nennung folgt der XML-Ausgabe in der OAI-PMH-Schnittstelle, siehe Anhang.

MODELINFOID

Wird zur eindeutigen Identifikation der gesamten Modellannotationen benutzt und dient als Basis zur Bildung einer http-URI für das Modell, vgl. RELATIONS.

⁷⁵ vgl. (Bekiari, Doerr, & Bœuf, 2012, S. 37): “This class comprises the intellectual or artistic realisations of works in the form of identifiable immaterial objects, such as texts, poems, jokes, musical or choreographic notations, movement pattern, sound pattern, images, multimedia objects, or any combination of such forms that have objectively recognisable structures. The substance of F2 Expression is signs. Expressions cannot exist without a physical carrier, but do not depend on a specific physical carrier and can exist on one or more carriers simultaneously.” Ein Alignment zu frbr:Work ist ggf. denkbar, vgl. Diskussion Work – Expression für intellektuelle Einheiten in der „EDM – FRBRoo Application Profile Task Force (EFAP-TF)“, <http://pro.europeana.eu/web/network/europeana-tech/-/wiki/Main/Task+Force+EDM+FRBRoo>, (oder Diskussion zu „Work“ beispielsweise unter <http://kcoyle.blogspot.de/2009/08/what-is-frbr-work.html>). Allerdings ist analog der in Kapitel 1 angesprochenen Modelltheorie die Klasse/ Entität „Expression“ die ausdrucksstärkste Variante für die Entität MODELINFO, da ein prä- oder deskriptives Modell einer Diskurswelt ebenso in anderen Formen (auf Papier oder mittels Bauklötzen oder mittels eines alternativen Modellierungstools) als intellektuelle oder künstlerische Realisierung eines Werks respektive imaginären Modells realisiert werden könnte.

Namespace:	dc:identifier
Occ.:	1

TITLE

Modelltitel

Namespace	dc:title
Beispiele	„Modell des Brandenburger Tors“
Occ.	1

MODELTYPE

Für die Entscheidung, welche Methoden zur Prozessierung und automatischen Erschließung angewendet werden, ist die Information über den vorliegenden Modelltyp (Objektmodell / Gebäudemodell / Umgebungsmodell / Szenenmodell) zwingend notwendig. Die hier aufgeführten Typen sind domänenspezifisch. Beispielsweise können Techniken der Extraktion struktureller Attribute wie Raumverbindungsgraphen nur auf Gebäudemodelle angewendet werden.

Namespace	dcterms:type P2 has type: E55 Type (P3 has note ⁷⁶)
-----------	---

⁷⁶ Die Adaption von “P3 has note” impliziert eine hier ggf. notwendige Flexibilität, vgl. (Bekiari u. a., 2012, S. 145): “Like property P2 has type (is type of), this property is a consequence of the restricted focus of the CRM. The aim is not to capture, in a structured form, everything that can be said about an item; indeed, the CRM formalism is not regarded as sufficient to express everything that can be said.”

Nicht nur an dieser Stelle wird das Fehlen eines Standardvokabulars für Typen als Problem identifiziert. Allerdings ist zu diskutieren, ob es eine solche Liste sowohl für domänenübergreifende als auch domänenspezifische Typen geben sollte (beispielsweise im EDM gibt es noch keine kontrollierten Listen für Eventtypen) oder die Anforderungen zu anwendungsspezifisch sind. Mit der fortschreitenden ontologischen Wissensmodellierung sind hierfür Antworten zu erwarten.

Beispiele	„Objektmodell“ „Gebäudemodell“
Occ.	1

REPRESENTATION

Ähnlich wie Modeltype, aber domänenunabhängig. Gibt Informationen über die geometrische Repräsentationsform, die sich z.T. automatisch aus den Dateiendungen ableiten lässt.

Namespace	dcterms:type P2 has type: E55 Type (P3 has note ⁷⁷)
Beispiele	„Oberflächenmodell“ „Volumenmodell“
Occ.	0-1

SUBJECTAREAS

Fachgebiet (Architektur / Bauingenieurwesen / Archäologie), denen das Modell zugeordnet werden kann, wobei eine Zuordnung zu mehreren Fachgebieten möglich ist.

Namespace	dcterms:type P2 has type: E55 Type (P3 has note ⁷⁸)
Beispiele	„Bauingenieurwesen“

⁷⁷ s.o.

⁷⁸ s.o.

Occ.	0-n
------	-----

MODELDESCRIPTION

Beschreibender Text zum Modell. Dieser kann von den Modellautoren / Bereitstellern bei Bedarf hinzugefügt werden.

Namespace	dc:description
Occ.	0-n

IDENTIFIER

Objektidentifikation, der zum Zweck des Downloads dient. Z.B. eine DOI, die zu einer URL aufgelöst werden kann. IDENTIFIER ist kein weiterer Identifier der Entität MODELINFO, sondern eine Beziehung zum Originalmodell.

Beispiele	„doi:10.5438/1234“
Occ.	0-n

ACCESSIBLE

Beinhaltet Informationen zur Downloadmöglichkeit der Modelldateien.

Namespace	P3 has note
Beispiele	„Accessible for download“
Occ.	1

LICENSE

Lizenzangaben zum Modell und seinen Dateien, die sowohl in Kurzform, ausführlicher Beschreibung als auch URI-Referenz vorliegen können, vgl. Umsetzung im Anhang.

Namespace	dc:rights / dcterms:rights
Beispiele	CC BY 3.0/de

	Namensnennung 3.0 Deutschland
Occ.	1

DATES

Für das 3D-Modell und seine Repräsentation im WWW relevante Datumsangaben. Ausführlich im Anhang dargelegt.

Namespace	dcterms:available, dcterms:created, dcterms:dateAccepted, dcterms:dateCopyrighted, dcterms:dateSubmitted, dcterms:modified
Beispiele	<DATEUPDATED>2012-03-01</DATEUPDATED> <DATEAVAILABLE>2008-01-01</DATEAVAILABLE>
Occ.	1-n

PREVIEWS

Modellrepräsentation zur Bereitstellung im WWW. Inkludiert die http-URI des Modells und 2D- / 3D-Vorschaudateien

Namespace	edm:landingPage → edm:WebResource ⁷⁹
Beispiele	„http://3d.probado.igd.fraunhofer.de/ Probado3D/ModelDetails.aspx?ID=205“
Occ.	1-n

MODELATTRIBUTES

„Medium-level“ Annotation. Zur Belegung unterschiedlicher domänenspezifischer Attribute, insbesondere maschinell erzeugter. Neue Attribute können als

⁷⁹ “This class is used as the type of the digital representations that are aggregated to the cultural heritage object. As such, it is the range of edm:landingPage”, vgl. (Europeana v1.0 Group, 2012, S. 15).

Einträge aufgenommen werden, ohne dass das Datenmodell erweitert werden muss.

Namespace	E70 Thing <i>P43 has dimension</i> : E54 Dimension ⁸⁰ <i>P2 has type</i> : E55 Type <i>P90 has value</i> : E60 Number <i>P91 has unit</i> : E58 Measurement Unit
Beispiele	Stockwerksanzahl, Nutzfläche, Raumanzahl, Geschossanzahl, umbauter Raum, ...
Occ.	0-n

Die unter E55 Type aufzuführenden Attribute können sich aus architekturenspezifischen Systemen wie AAT (wenn zukünftig als Linked Data verfügbar) speisen. Anzahl und Aussagekraft domänenspezifischer Ontologien ist jedoch ausbaubar und wird ggf. durch im Kontext BIM / IFC stattfindende Entwicklungen wie das International Framework of Dictionaries (IFD)⁸¹ angegangen.

MODELINFO → MODELINFO / a. Dokumente (RELATIONS)

Zur Identifikation von Modellen und anderen Dokumenttypen, die in Relation zum Ausgangsmodell stehen.

Hier werden Aggregationen von unterschiedlichen Dokumenten unterstützt und sowohl hierarchische als auch Assoziativbeziehungen ermöglicht.

Namespace	dc:relation
Occ.	0-n

⁸⁰ MODELINFO (F2 Expression) ist Sub-class von E73 Information Object und damit Sub-class von E70 Thing.

⁸¹ vgl. http://www.ifd-library.org/index.php?title=Technical_Documentation bzw. http://dev.ifd-library.org/index.php/Ifd:IFD_API#Classes

RELATIONROLE

MODELINFO-Beschreibungen werden mittels Rollen zueinander und zu anderen Dokumenttypen in Relation gesetzt. Z.B. zur Definition einer „IsSupplementTo“-Beziehung eines Modells zu einem Textdokument oder zur Definition einer rekursiven „IsPartOf“-Relation eines geschachtelten Modells.

Namespace	dcterms:... (siehe Auflistung im Anhang)
-----------	--

Außerdem kann MODELINFO Relationen zu weiteren, in den folgenden Abschnitten ausführlich behandelten Entitäten beinhalten:

MODELINFO → MODELFILE

Instanzen eines 3D-Modells. Meistens liegt ein Modell in unterschiedlichen Formaten vor, die behandelt werden können. Somit kann MODELINFO mehrfache Relationen zum MODELFILE beinhalten. An dieser Stelle werden nur die primären CAD-Dateien angeführt und nicht z.B. Texturen.

Namespace:	frbroo:R6i_is_carried_by ⁸²
Occ.	1-n

MODELINFO → CONTRIBUTOR

Akteure, siehe 4.5.1, die in unterschiedlichen Rollen, u.a. mit 3D-Modellen assoziiert werden können.

Namespace	dcterms-Properties ⁸³ , spezifische Rollen s.u.
-----------	--

⁸² Es gibt noch keinen offiziellen Namespace. In FRBR fehlte noch eine Property zwischen der abstrakten Entität und dem konkreten Item – mit Domain Work/ Manifestation und Range Item. In FRBRoo wird diese Relation mittels [R6 carries (is carried by), Domain: F5 Item, Range: F24 Publication Expression.] hergestellt, vgl. (Bekiari u. a., 2012, S. 63).

⁸³ Eine entsprechende Property gibt es in FRBRoo nicht, da CRM an dieser Stelle einen ereignisorientierten Ansatz wählt. D.h., es kann gar keine direkte Beziehung zwischen MODELINFO und CONTRIBUTOR in FRBRoo oder CRM geben, denn sie ist immer ereignisvermittelt. Es wird daher auf DCterms zurückgegriffen. Eine exakte Modellierung dieser Relation auch in FRBRoo ist wünschenswert. -- Erste Überlegung war die Anwendung von “isRealizedBy”: “the “realized by”

Occ.	1-n
------	-----

ROLE

Contributors werden mit MODELINFO mittels Rollen in Relation gesetzt (als Ersteller, Eigentümer, Urheber, ..)

Namespace	dc:creator: responsible for creation of model dc:publisher: responsible for accessibility of model crm:right held by (has right on) dc:Contributor
Beispiele	Creator Publisher / Distributor Rightsholder Related Contributor

MODELINFO → SUBJECT

beinhaltet Konzepte zur inhaltlichen Beschreibung von 3D-Modellen. Auch die automatisch erzeugten Klassifikationen werden hier abgelegt, wobei unterschieden wird, ob manuell oder maschinell klassifiziert wurde, vgl. 4.6.4

Namespace:	dcterms-Properties, spezifische Rollen siehe 4.6.4
Occ.	0-n

relationship that links the same two entities to expression. [...] It is similar in function to the “created by” relationship, but it implies a difference in the nature of the relationship that parallels the difference between work and expression as entities.”, vgl. (IFLA Study Group on the Functional Requirements for Bibliographic Records & International Federation of Library Associations and Institutions, 1998, S. 61;64).

MODELINFO → EVENT

beschreibt Ereignisse, die zur Erstellung des Modells beigetragen haben können wie beispielsweise Wettbewerbe, Workshops, Ausstellungen oder Seminare.

Namespace:	P19i_was_made_for ⁸⁴
Occ.	0-n

MODELINFO → MODOBJECT

beinhaltet Informationen zu Objekten der realen Welt wie Gebäuden oder Gegenständen.

Namespace:	P138 represents
Occ.	0-n

4.4.2 MODELFILE

Die Entität MODELFILE stellt digitale Instanzen einer 3D-Modellierung bereit und inkludiert alle Angaben bzgl. technischer Metadaten einer in einem bestimmten Format erzeugten Datei, z.B. Dateigröße.

Namespace:	F5 Item ⁸⁵
------------	-----------------------

⁸⁴ Die Relation gibt es so nicht in CRM. Im EDM: Event edm:happenedAt Place. Von der Definition in FRBRoo her passt ggf. P19i_was_made_for.

⁸⁵ (Bekiari u. a., 2012, S. 40) definieren frbr:Item als materielles Objekt: „This class comprises physical objects (printed books, scores, CDs, DVDs, CD-ROMS, etc.) that carry a F24 Publication Expression“. -- Unter Berücksichtigung der in Kapitel 1 beschriebenen Ausgabe als 2D- oder 3D-Druck, wäre zu diskutieren, ob diesen konkreten, materiellen Objekte nicht der Namespace „Item“ zugeschrieben werden sollte. Zwar werden haptische Ausgaben von MODELIFO nicht im Modell selbst berücksichtigt, aber eine Abgrenzung zwischen digitalen und materiellen Objekten als Items, vgl. auch Übersicht unter http://www.cidoc-crm.org/docs/EDM-DC-ORE-CRM-FRBR_Integration_ORE_fix.ppt, S. 2, scheint hier sinnvoll und notwendig. Daher kann, pragmatisch das Items als Dokument bzw. Ressource betrachtend, alternativ zu frbr:Item aus der Bibliographic Ontology <http://bibliontology.com/> die Ressource bibo:AudioVisualDocument als Namespace dienen.

Die meisten Attribute aus MODELFILE eint, dass sie während der in 3.4 dargestellten Prozessierung zur Informationsextraktion und Vorbereitung für das inhaltsbasierte Retrieval in einem frühen Schritt der Kette, vgl. Abbildung 12, automatisiert erzeugt werden können. Da das Anbringen von Beispielen meist nicht notwendig zum Verständnis scheint, werden sie hier in kürzerer Form und teilweise zusammengefasst aufgeführt.

FILENAME

Dateiname ohne Dateiendung.

Namespace	dc:title
Occ.	1

FILESIZE

Eine wichtige Information u.a. für den Download.

Namespace	rda:FileSize
Occ.	1

FORMAT

Format / Dateiendung. Siehe auch Darstellung im Anhang.

Namespace	dc:format rda:EncodingFormat
Occ.	1

FILEATTRIBUTES

„Low-level“ Annotation. Zur Belegung weiterer, insbesondere maschinell extrahierbarer Attribute, wie die Menge von Knotenpunkten und Polygonen. Solche

Informationen sind für den Benutzer von Interesse zur Einschätzung des Detaillierungsgrades⁸⁶ der Modelldatei.

Namespaces	P43 has dimension ⁸⁷ <i>P2 has type:</i> E55 Type <i>P90 has value:</i> E60 Number <i>P91 has unit:</i> E58 Measurement Unit rda:DigitalFileCharacteristic
Occ.	0-n

Insbesondere das Dateiformat ist dafür ausschlaggebend, ob eine Datei für die Integration verwendet werden kann und dient damit als wichtige Annotation zum Finden passender Modelldateien, vgl. Kapitel 1. Die Liste der Attribute ist prinzipiell entsprechend der aus den Dateien maschinell extrahierbaren Informationen beliebig erweiterbar. Bzgl. Namespaces wird über Dimensions derselbe Weg gegangen wie bei MODELINFO:MODELATTRIBUTES. Alternativen sind direkte Ontologiemappings, z.B. Höhenkoordinate - yago:Axis106008609, Units - geoR⁸⁸:hasSpatialUnits (Range: Millimeters, Centimeters, Meters), Vertices – OpenCyc⁸⁹:numberOfVertices.

Das Attribut Units (Zeicheneinheiten) sowie Höhenkoordinate sind Voraussetzung für die Adaption diverser Verfahren der inhaltsbasierten Erschließung, insbesondere zur Erzeugung der in (Wessel, Blümel, & Klein, 2008) eingeführten Raumverbindungsgraphen. Diese Attribute sind nicht maschinell extrahierbar, jedoch lassen die Dateiendungen auf die verwendete Software und damit auf den defaultmäßig eingestellten Zeichenmaßstab und die Höhenkoordinate schließen.

⁸⁶ Auch Trimble / Google stellt in seinem bereits angesprochenen Data Warehouse solche Informationen bereit.

⁸⁷ vgl. MODELATTRIBUTES

⁸⁸ <http://www.mindswap.org/2003/owl/geo/geoRelations20040307.owl#>

⁸⁹ <http://sw.opencyc.org/>

SOFTWARE

Beinhaltet Informationen über die angewendete CAD- / Modellier-Software. SOFTWARENAME: Software Bezeichnung (skos:prefLabel), SOFTWARE-VERSION. Zusätzliche Verbindung zu CORPORATE : Software Firma möglich. Siehe Darstellung im Anhang.

Namespace	yago:Software / Version, Firma sind einzuführen
Beispiele	„AutoCAD“ / „Version 12“ / „Autodesk“

Das Attribut SOFTWARE kann nur bei proprietären Dateiformaten und mit Vorhandensein maschinell auswertbarer Zuordnungen (Format – Software – Firma usf.) automatisiert belegt werden.

4.5 Gruppe 2 – involvierte Personen und Körperschaften

Gruppe 2 umfasst Informationen zu Personen oder Körperschaften, die mit 3D-Modellen, mit MODOBJECT oder mit EVENT in Beziehung stehen.

4.5.1 CONTRIBUTOR

assoziierte Akteure

Namespace	E39 Actor
-----------	-----------

MODELINFO → CONTRIBUTOR

Akteure, die in unterschiedlichen Rollen u.a. mit 3D-Modellen assoziiert werden können, vgl. Darstellung unter 4.4.1.

CONTRIBUTOR → MODOBJECT

Architekt, Designer

Namespace:	E39 Actor P14i_performed E65 Creation P94_has_created E28 Conceptual Object ⁹⁰
Occ.	0-n

CONTRIBUTOR → EVENT

Eventausrichter (als wettbewerbsauslobende Institution, Institut des Workshops, Seminars o.ä.)

Namespace:	crm:performed (carried out by)
Occ.	0-n

Ein Contributor kann entweder eine Person oder Körperschaft sein, wobei Personen und Körperschaften miteinander in Relation gesetzt werden können.

4.5.2 PERSON

Natürliche Personen

Namespace	(foaf:Person) F10 Person / E21 Person
-----------	--

LASTNAME

Namespace	foaf:familyName
-----------	-----------------

FIRSTNAME

Namespace	foaf:firstName
-----------	----------------

EMAIL

Notwendig, um beispielsweise den studentischen Modellersteller kontaktieren zu können.

⁹⁰ “has created (was created by)” reicht im CRM nicht aus. Es wird unter Einbeziehung eines Events modelliert.

Namespace	pim ⁹¹ :EmailAddress
-----------	---------------------------------

EXTERNALINFO

URI als eindeutige personenbezogene Information. Z.B. aus DBPedia, Personen-normdatei.

Namespace	rdfs:about edm:NonInformationResource
Occ.	0-n

PERSON → CORPORATE

Beziehung zwischen Person und Institution

Namespace:	ecs:memberOf ⁹²
Occ.	0-n

4.5.3 CORPORATE

Körperschaften

Namespace	foaf:Organization F11 Corporate Body
-----------	---

NAME

Namespace	skos:PrefLabel
Occ.	1

⁹¹ foaf bietet keine Property. Daher http://www.w3.org/2000/10/swap/pim/contact#_EmailAddress

⁹² vgl. <http://rdf.ecs.soton.ac.uk/ontology/ecs>. Foaf bietet analog foaf:membershipClass. Beides passt allerdings nicht optimal.

ADRESS

Wird ausführlich im Anhang dargestellt.

Ortsangaben unter Person und Corporate könnten auch unter Place zusammengefasst werden. Hier als personen- / firmenbezogene Daten eher als Adresse gehandhabt.

Occ.	0-n
------	-----

EXTERNALINFO

URI als eindeutige Information.

Namespace	rdfs:about edm:NonInformationResource
Occ.	0-n

PARENT (CORPORATE → CORPORATE)

Link zur übergeordneten Körperschaft, z.B. Institut → Fakultät → Universität. Notwendig, da Modelle von unterschiedlichen Einheiten / Ebenen derselben Institution bereitgestellt werden.

Namespace:	ecs:affiliatedWith
Occ.	0-n

4.6 Gruppe 3 –Ereignisse, reale Objekte, Orte und Konzepte

Gruppe 3 enthält Entitäten, die zusätzliche Informationen über ein Modell liefern können, und zwar mittels der Entitäten EVENT (Ereignis, z.B. Wettbewerb, Workshop, Ausstellung), MODOBJECT (gebaute Architektur, Konstruktion oder Designobjekt), PLACE (Orte von Events und ModObjects) und SUBJECT (Konzept, Begriff, Klassifikation).

4.6.1 EVENT

Ereignisse, die zur Erstellung des Modells beigetragen haben. Über EVENT lassen sich Modelle zu Gruppen bzw. Sammlungen aggregieren⁹³.

Namespace	F8 Event F28 Expression Creation ⁹⁴
-----------	---

EVENTTYPE

Dient zur Typisierung und Gruppierung von Ereignissen. Beispielsweise eines Seminars oder Wettbewerbs; auch einer Sammlung.

Namespace	dcterms:type P2 has type E55 Type
Occ.	1

EVENTDATE

Zeitangabe zum Event.

Namespace	crm:P4_has_time-span
Occ.	0-n

⁹³ Aggregationen ließen sich auch mittels Nachnutzung des OAI Object Reuse and Exchange (ORE) Namespace (<http://www.openarchives.org/ore>) erreichen, beispielsweise angewendet als „Proxies“ und „Aggregations“ im EDM. Dadurch dass Modellschachtelungen in MODELINFO abgebildet werden können, ist zusammen mit der Aggregation über EVENT zum Zeitpunkt der Implementation des Datenmodells noch keine Notwendigkeit für elaboriertere Modellierungen dieser Konstrukte gesehen worden.

⁹⁴ Für sonstige Aktivitäten, die zur Erstellung (und Publikation → Kindklasse „F30 Publication Event“ adaptieren) von Modellen beitragen, passt diese Klasse genauer als „F8 Event“. Vgl. (Bekiari u. a., 2012, S. 51) “This class comprises activities that result in instances of F2 Expression coming into existence.”

EVENTDESCRIPTION

Beschreibender Text zum Event, beispielsweise die Wettbewerbskonditionen oder Aufgabenstellung bzw. Definition der Entwurfsabgabeleistung eines Seminars.

Namespace	dc:description
Occ.	0-n

EXTERNALINFO

URI als eindeutige http-Adresse zur Identifikation des Events.

Namespace	rdfs:about edm:NonInformationResource
Occ.	0-n

EVENT → CONTRIBUTOR

Akteure, die mit Events assoziiert werden können, z.B. der Eventauslober, Institut des Seminars.

Namespace	crm:P11_had_participant
Occ.	0-n

EVENT → PLACE

Orte, die mit Events assoziiert werden können, z.B. Ort der Konferenz oder Bauausstellung / Ausstellung der Wettbewerbsergebnisse.

Namespace	dcterms:spatial P7 took place at
Occ.	0-n

4.6.2 OBJECT

Diese Entität umfasst reale (fallweise auch imaginäre Objekte, vgl. Abschnitt 2.1.2) beispielsweise bereits existierende Gebäude, Konstruktionen oder Design-

objekte wie Möbel oder dreidimensionale Artefakte z.B. aus dem Bereich des Kulturerbes. Reale Objekte sind neben Konzepten (vgl. SUBJECT) die wichtigsten Informationsträger auf der Suche nach 3D-Modellen, insbesondere wenn sie deskriptiven Charakter haben, vgl. Übersicht unter 2.1.2.

Namespace	F7 Object E18 Physical Thing
-----------	---------------------------------

P53 has former or current location (is former or current location of): E53 Place

Bei der Auswahl relevanter Attribute und Relationen wurden Nutzeranforderungen einbezogen (vgl. Szenarien in 2.4.2), also die bei der Suche nach einem 3D-Modell möglichen Parameter, z.B. der Ort oder Architekt eines realen Bauwerks. Die Attribute beschränken sich im Wesentlichen auf Bezeichnungen sowie auf Relationen zum Designer / Architekten / Ersteller (CONTRIBUTOR) des realen Objekts, zu dessen dreidimensionaler Repräsentation, dem 3D-Modell (MODEL-INFO) und zum Ort (PLACE), falls es sich, wie bei Gebäuden der Fall, um ein verortetes Objekt handelt. Die Anbindung zusätzlicher Informationen⁹⁵ (z.B. Events, die zur Erstellung des realen Objekts beigetragen haben, oder Zeitangaben wie Errichtung, Umbau, Abriss, usw.) wird mittels URIs ermöglicht, sodass die Entität selbst relativ kompakt gehalten werden kann.

OBJECTNAME

Benennung des realen Objekts.

Namespace	dc:title P1_is_identified_by
-----------	---------------------------------

⁹⁵ (Boeykens & Bogani, 2008) merken dazu an: “There is often no obvious difference between metadata describing a building project and metadata from a digital model of that building. The values for fields such as designer, location, date (timeframe), context, building type, style, construction and others pertain to both the real world object and the 3D object [...]”. Dies beschreibt die Problematik bei der Entscheidung, welche Metadaten extern liegen und welche zwingend im Modell abgebildet werden müssen. So sind z.B. detaillierte Informationen zur Konstruktion sinnvoller am bzw. besser durch das 3D-Modell zu liefern, als sie am realen Objekt abzulegen.

Occ.	1
------	---

OBJECTDATE

Zeitangabe bzgl. Fertigstellung des realen Objekts.

Namespace	dc:date
Occ.	0-n

EXTERNALINFO

URI als eindeutige http-Adresse zur Identifikation des realen Objekts, z.B. aus der DBPedia.

Namespace	rdfs:about edm:NonInformationResource
Occ.	0-n

OBJECT → MODELINFO

Verbindung zur dreidimensionalen Repräsentation des realen Objekts, siehe Darstellung unter 4.4.1.

OBJECT → PLACE

Orte, die mit realen Objekten assoziiert werden können, z.B. Ort des Bauwerks. Durch die Integration der URI kann diese Information auch an externer Stelle abgefragt bzw., soweit sie nach Linked Data Prinzipien strukturiert ist, importiert werden.

Namespace	dcterms:spatial
Occ.	0-n

OBJECT → CONTRIBUTOR

Akteure (Personen und Körperschaften), die als Ersteller derselben mit realen Objekten assoziiert werden können, soweit sie bekannt sind, z.B. Architekten, Desig-

ner, Architekturbüros, siehe Darstellung unter 4.5.1. Durch die Integration der URI kann diese Relation auch durch Reasoning über angebundene Daten erzeugt werden, soweit die Daten nach Linked Data Prinzipien strukturiert sind.

4.6.3 PLACE

Mit Events oder Objects assoziierte Orte. Ortsangaben bieten einen möglichen Einstieg bei der Suche. Ein Beispiel im Rückgriff auf die Szenarien aus Kapitel 1 wäre, dass alle Modelle realer Gebäude einer bestimmten Stadt gesucht werden.

Namespace	frbr:Place
-----------	------------

PLACENAME

Benennung des Ortes.

Namespace	dc:title P1_is_identified_by
Occ.	1

LAT

Latitude. GK-Koordinate zur genauen Positionsbestimmung, z.B. Verortung auf einer Karte.

Namespace	geo ⁹⁶ :lat
Occ.	0-1

LON

Longitude. GK-Koordinate zur genauen Positionsbestimmung, z.B. Verortung auf einer Karte.

Namespace	geo ⁹⁷ :long
-----------	-------------------------

⁹⁶ GeoNames Ontology, vgl. http://www.w3.org/2003/01/geo/wgs84_pos#

Occ.	0-1
------	-----

ADDRESS

Hier können Adressinformationen eines Ortes angefügt werden. Details siehe Anhang.

Occ.	0-1
------	-----

EXTERNALINFO

URI als eindeutige http-Adresse zur Identifikation des Ortes. Z.B. aus der DBpedia, Geonames oder, wenn zukünftig als LOD vorhanden⁹⁸, dem geographischen Thesaurus des Getty-Instituts.

Namespace	rdfs:about edm:NonInformationResource
Occ.	0-n

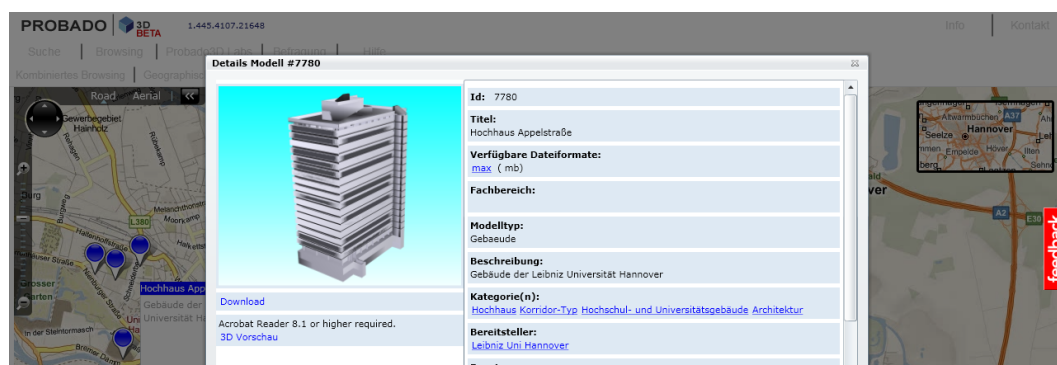


Abbildung 15: 3D-Gebäudemodell eines real existierenden Bauwerks mit dargestellter Verortung (Quelle: PROBADO 3D)

⁹⁷ s.o.

⁹⁸ http://www.getty.edu/research/tools/vocabularies/Linked_Data_Getty_Vocabularies.pdf

4.6.4 SUBJECT (CONCEPT)

„High-level“ Attribute. Die Entität SUBJECT beinhaltet Konzepte und Klassifikationen, die MODELINFO zugeordnet werden können. Sehr wichtig als Sucheinstieg, beispielsweise aller Gebäude, die mit einem bestimmten Baustil gelabelt werden. Gleichmaßen für die Suche über Funktionen und Typen wie „Krankenhausbau“, vgl. (Blümel u.a., 2010).

Namespace	F6 Concept
-----------	------------

DEFAULTNAME

Standardbenennung des Konzepts.

Namespace	skos:prefLabel
Occ.	1

Durch SUBJECT sollen die 3D-Modelle mit bestehenden Vokabularen, z.B. bibliothekarischen Normdaten, verschlagwortet bzw. beschrieben werden. Notwendig ist die Bereitstellung von Links in den Metadaten zu diesen Vokabularen. Dazu dienen die Attribute NOTATION oder – im Hinblick auf die maschinelle Interpretationsfähigkeit besser anwendbar – EXTERNALINFO.

NOTATION

Dient zur Integration von Notationen aus Thesauri, Ontologien oder Klassifikationen zur Identifikation des jeweiligen Schlagworts / Konzepts.

Namespace	skos:notation
Beispiele	„725.2“ (DDC-Notation)
Occ.	0-n

SUBJECTSCHEME

Name des jeweiligen Begriffssystems.

Namespace	skos:inScheme
-----------	---------------

Beispiele	“Getty Art and Architecture Thesaurus (AAT)” “Dewey Decimal Classification (DDC)”
Occ.	1

Da immer mehr bibliothekarische Normdaten als Linked Open Data zur Verfügung gestellt werden, kann mittelfristig das dargestellte Konstrukt „NOTATION:SUBJECTSCHEME“ durch eine URI des jeweiligen Begriffs ersetzt werden, vgl. folgende Property.

EXTERNALINFO

URI als eindeutige http-Adresse zur Identifikation des Begriffs, z.B. aus der GND.

Namespace	rdfs:about edm:NonInformationResource
Occ.	0-n

SUBJECT→MODELINFO

Dient zur Verknüpfung der Begriffe mit den 3D-Modellen, siehe genauere Beschreibung unter 4.4.1.

CLASSIFICATION

Modus der Klassifizierung. Die Relation eines Konzepts zu MODELINFO muss attribuiert werden, d.h. es muss Auskunft darüber gegeben werden, wie die Klassifikation erfolgt ist: manuell, mittels Crowdsourcing, maschinell. Diese Information ist insbesondere wichtig hinsichtlich der Verlässlichkeit der Annotation.

Beispiele	manuell maschinell
Occ.	1

CONFIDENCE

Konfidenzwert der Klassifikation.

Beispiele	<p>„0,85“</p> <p>Bei manueller Klassifikation wird als Konfidenzwert „1“ eingetragen.</p>
Occ.	1

Maschinell erzeugte Klassifikationsergebnisse können probabilistischer Natur sein, d.h. ein Klassenlabel wird dann mit einer gewissen Konfidenz vergeben (bzw. nicht vergeben), vgl. Abschnitt zur maschinellen Erschließung → Autoklassifikation in 3.4.2. Bei der Suche nach 3D-Modellen kann diese Konfidenz berücksichtigt werden, sodass auf den vorderen Ergebnisseiten nur solche Modelle angezeigt werden, bei denen das jeweilige Klassenlabel relativ sicher vergeben wurde, siehe Abbildung 16. Die Modelle mit zweifelhafterer Labelvergabe, bei denen erfahrungsgemäß die false positive Rate höher ist, tauchen erst auf den hinteren Ergebnisseiten auf.

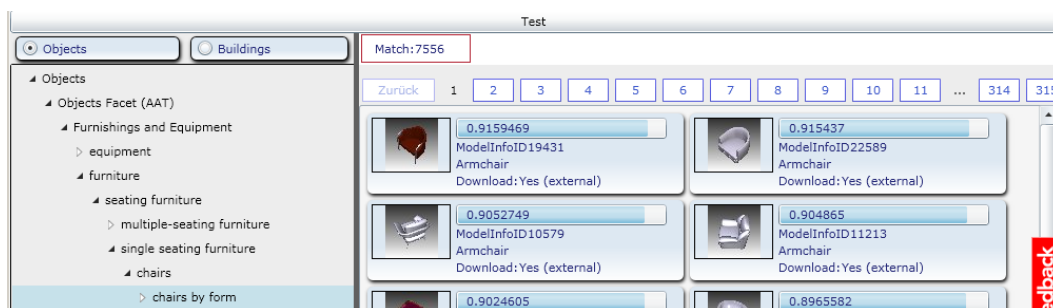


Abbildung 16: Browsing über 3D-Modelle mittels Begriffen, die durch automatisierte Klassifikation vergeben wurden. Anzeige der Konfidenz der Klassifikation am jeweiligen Modell (PROBADO 3D)

4.7 Visualisierung

Die Darstellung in Abbildung 17 zeigt eine graphen- bzw. aussagenbasierte Visualisierung der entsprechend des beschriebenen Metadatenmodells strukturierten Informationen zu einem exemplarischen 3D-Modell und dessen Kontext. Avancierte Fragen der graphen- / RDF-basierten Modellierung, z.B. Themen, wie Provenienz- und Versionierungsinformationen über Metadaten, werden hier nicht

behandelt. Da Linked Data nur den zuletzt publizierten Zustand von Ressourcen und die Verbindungen unter ihnen kennt, müssen umständliche, wenig performante Meta-Aussagen helfen. Eine momentan diskutierte Erweiterung des RDF-Modells sind sogenannte „Named Graphs“⁹⁹. Diese können auch als Lösung zur Referenz auf das gesamte Metadatenkonstrukt dienen. Zur Beschreibung von Metadatenkonstrukten können ebenfalls, wie im EDM angewendet, Proxies und Aggregationen mittels des OAI Object Reuse and Exchange (ORE)¹⁰⁰ Namespace herangezogen werden.

⁹⁹ <http://www.w3.org/2004/03/trix/>

¹⁰⁰ <http://www.openarchives.org/ore>

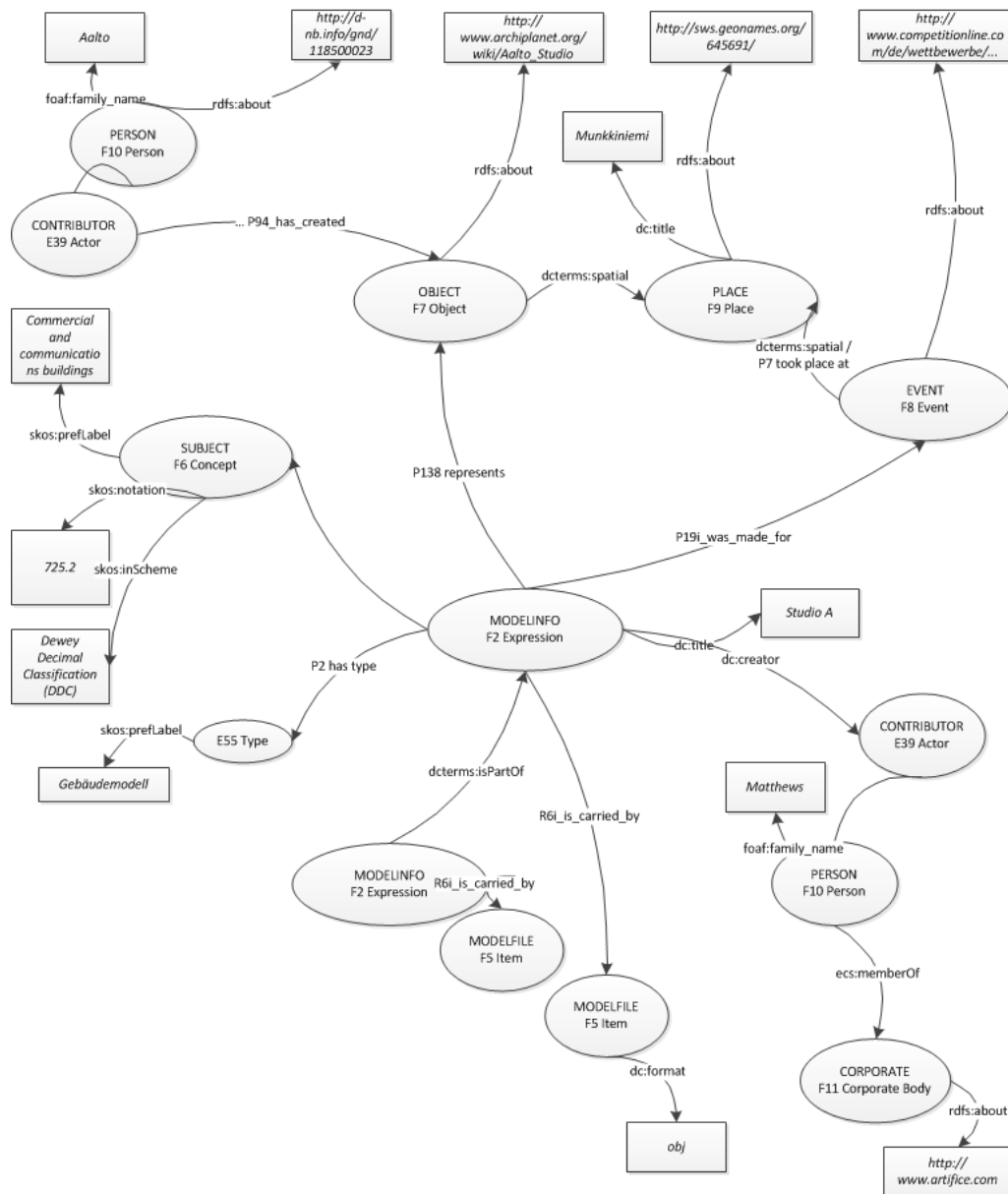


Abbildung 17: Aussagen über ein 3D-Modell und dessen Kontext

Abbildung 18 visualisiert die zu einem beispielhaften 3D-Modell vorhandenen Metadaten (entsprechend der ab Abschnitt 4.3 gelisteten Entitäten). Es handelt sich aus Gründen der Übersichtlichkeit hierbei nicht um eine vollständige Auflistung aller potenziellen Informationseinheiten; insbesondere hinsichtlich der in Kapitel 1 erörterten automatisch extrahierbaren inhaltlich-technischen Metadaten wären noch weitere hinzuzufügen. Die Darstellung zeigt ferner vorhandene Kontextinformationen.

99

ten von Linked Data und hinsichtlich der Einbindung von Kontextinformationen.¹⁰¹

¹⁰¹ vgl. <http://drunks-and-lampposts.com/2012/06/13/graphing-the-history-of-philosophy/>,
<http://griffsgraphs.com/>, <http://delicious.com/noelmas/Humanities-Visualisations>,
<http://delicious.com/noelmas/open-cultural-heritage-resources>

5 Bewertung und Ausblick

Ziel der vorliegenden Arbeit war es, ein Metadatenmodell zu entwickeln, das diskrete Informationseinheiten in und um 3D abbilden kann, unabhängig davon, ob die Erschließung maschinell oder manuell oder in Kombination beider Verfahren stattfindet.

Zu diesem Zweck wurden inhaltliche und technische Anforderungen aus der 3D-Architekturdomäne formuliert sowie aktuelle Möglichkeiten der Wissensorganisation und automatisierter Informationsextraktion untersucht. Dabei ergab sich, dass die aussagenbasierte, strukturierte Annotation analog der Prinzipien des Linked Data eine zukunftsweisende Technologie darstellt, 3D-Modell-strukturspezifische und technische sowie inhaltlich-kontextuelle Information flexibel abzubilden. Diese Herangehensweise hat im Gegensatz zu einem Ansatz, der Metadaten als Records auffasst, das Potenzial, Relationen zu jeglichen modellrelevanten Entitäten im Suchraum herzustellen und zugleich diese Relationen wieder für weitere wissensbildende Prozesse verfügbar zu machen. Am Beispiel von Gebäude- und Objektmodelldaten wurde gezeigt, wie diese Information strukturiert werden könnte bzw. sollte.

Der pragmatische Ansatz, Namensräume etablierter, generischer Modelle und Ontologien bei möglichst konsistentem Alignment zu adaptieren, um ein hohes Maß an Interoperabilität und Nutzbarmachung der Modellbeschreibungen als Linked Data zu gewährleisten, lässt eine relativ umfassende Abbildung der in Kapitel 2 und 3 als 3D-domänenspezifisch relevant erörterten Aussagen zu. Während für domänenunabhängige und mittels automatischer Extraktionsmethoden gewonnene technische 3D-Modellmetadaten bereits ontologische Typ-Entsprechungen gefunden werden konnten, bedarf es für die volle Ausdrucksstärke automatisiert erstellter inhaltlich-technischer Metadaten weiterer, domänenspezifischer Ontologien.

Zwar kann anhand dieses Befundes darauf geschlossen werden, dass das Modell hinsichtlich automatisch extrahierter Metadaten unter Anwendung entsprechender Ontologien auf andere Fachbereiche übertragbar ist, jedoch gilt dies, wie in Abschnitt 3.4.2 gezeigt werden konnte, nur unter der Einschränkung, dass maschinell-

le Lernverfahren an die jeweilige Domäne angepasst werden. Außerdem ist zu überprüfen, ob neben den in drei Gruppen zusammengefassten Hauptentitäten in anderen Domänen ggf. weitere Konzepte hinzugezogen werden müssten.

Insgesamt kann festgestellt werden, dass das Modell – als Spezialisierung von FRBRoo und CRM für 3D-Modell-Metadaten – ein hohes Maß an Möglichkeiten zur Adaption gewährleistet und die – durch Aussagen- und Ontologien hergestellte – Flexibilität einen hinreichenden Rahmen zur Organisation der unterschiedlichen, insbesondere mittels automatischer Erschließung gewonnenen, Metadatentypen bietet. Insofern steht zu hoffen, dass sich in Zukunft die Ansätze für Ontologieentwicklung im Baubereich ausbauen lassen, um zusammen mit dem dargestellten Modell als ausdrucksstarke Beschreibungsinstrumente für 3D-Modelle zu dienen.

Literaturverzeichnis

- Aune, B. (1970). Rationalism, Empiricism, and Pragmatism: An Introduction. Abgerufen von <http://philpapers.org/rec/AUNREA>
- Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Information Review*, 13(5), 407–424. doi:10.1108/eb024320
- Beaudoin, J. E. (2012a). A Framework for Contextual Metadata Used in the Digital Preservation of Cultural Objects. *D-Lib Magazine*, 18(11/12). doi:10.1045/november2012-beaudoin2
- Beaudoin, J. E. (2012b). Context and Its Role in the Digital Preservation of Cultural Objects. *D-Lib Magazine*, 18(11/12). doi:10.1045/november2012-beaudoin1
- Bekiari, C., Doerr, M., & Bœuf, P. L. (2012, November 18). FRBR - object-oriented definition and mapping to FRBRER (version 2.0). Abgerufen 27. Dezember 2012, von http://www.cidoc-crm.org/docs/frbr_oo/frbr_docs/FRBRoo_V2.0_draft.doc
- Bernd, R., Blümel, I., Krottmaier, H., Wessel, R., & Schreck, T. (2009). Demonstration of User Interfaces for Querying in 3D Architectural Content in PROBADO3D. In *Proceedings of the 13th European Conference on Research and Advanced Technology for Digital Libraries* (S. 491–492). Berlin, Heidelberg: Springer-Verlag. Abgerufen von http://dx.doi.org/10.1007/978-3-642-04346-8_70
- Berndt, R., Blümel, I., Clausen, M., Damm, D., Diet, J., Fellner, D., ... Wessel, R. (2010). The PROBADO Project - Approach and Lessons Learned in Building a Digital Library System for Heterogeneous Non-textual Documents. In M. Lalmas, J. Jose, A. Rauber, F. Sebastiani, & I. Frommholz (Hrsg.), *Research and Advanced Technology for Digital Libraries* (S. 376–383). Springer Berlin Heidelberg. Abgerufen von http://link.springer.com/chapter/10.1007/978-3-642-15464-5_37, doi: 10.1007/978-3-642-15464-5_37
- Berndt, R., Blümel, I., & Wessel, R. (2010). PROBADO3D – Towards an Automatic Multimedia Indexing Workflow for Architectural 3D Models. In *ELPUB2010. Proceedings of 14th International Conference on Electronic Publishing 16-18 June 2010, Helsinki, Finland/ Edited by: Turid Hedlund and Yasar Tonta* (S. 79–88). Abgerufen von http://elpub.scix.net/cgi-bin/works/Show?_id=107_elpub2010

- Berners-Lee, T. (1997). Web architecture: Metadata. Abgerufen 8. Dezember 2012, von <http://www.w3.org/DesignIssues/Metadata.html>
- Berners-Lee, T. (2006a, April 5). *The Future of the Web*. Gehalten auf der Spencer Trask Lecture, Princeton University. Abgerufen von <http://www.w3.org/2006/Talks/0405-princeton-tbl/>
- Berners-Lee, T. (2006b, Juli 27). Linked Data - Design Issues. Abgerufen 27. Oktober 2012, von <http://www.w3.org/DesignIssues/LinkedData.html>
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web. A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*, 284, No. 5. Abgerufen von <http://www.scientificamerican.com/article.cfm?id=the-semantic-web>
- Bertram, J. (2005). *Einführung in die inhaltliche Erschließung: Grundlagen - Methoden - Instrumente* (1., Aufl.). Ergon.
- Bienert, A. (2012). *Elektronische Medien & Kunst, Kultur, Historie* □: EVA 2012 Berlin (1. Aufl.). Berlin: Gesellschaft zur Förderung angewandter Informatik.
- BLK. (2006). *Neuausrichtung der öffentlich geförderten Informationseinrichtungen: Abschlussbericht*. Bonn: Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung.
- Blümel, I., Berndt, R., Ochmann, S., Vock, R., & Wessel, R. (2010). Supporting Planning through Content-Based Indexing and 3D Shape Retrieval. *Proceedings of 10th International Conference on Design & Decision Support Systems in Architecture and Urban Planning*. Abgerufen von http://cg.cs.uni-bonn.de/aigaion2root/attachments/Indexing_and_Searching_3D_CAD_Databases.pdf
- Blümel, I., & Sens, I. (2009). Das PROBADO-Projekt: Integration von nichttextuellen Dokumenten am Beispiel von 3D-Objekten in das Dienstleistungsangebot von Bibliotheken. *Zeitschrift für Bibliothekswesen und Bibliographie*, 2/2009, 079–087.
- Boeykens, S., & Bogani, E. (2008, November). Metadata for 3D Models. How to search in 3D Model repositories? Abgerufen 12. Dezember 2012, von <https://lirias.kuleuven.be/handle/123456789/202356>
- Bustos Cárdenas, B. E., Keim, D. A., Saupe, D., Schreck, T., & Tatu, A. (2008). Methoden und Benutzerschnittstellen für effektives Retrieval in 3D-Datenbanken. article. Abgerufen 4. Dezember 2012, von <http://kops.ub.uni-konstanz.de/handle/urn:nbn:de:bsz:352-opus-68564>

- Candela, L., G. Athanasopoulos, D. Castelli, K. El Raheb, P. Innocenti, Y. Ioannidis, ... S. Ross. (2011, April). The Digital Library Reference Model. Abgerufen von <http://bscw.research-infrastructures.eu/pub/bscw.cgi/d222816/D3.2b%20Digital%20Library%20Reference%20Model.pdf>
- Caplan, P. (2003). *Metadata Fundamentals: For All Librarians*. ALA Editions.
- Capurro, R. (2006). Hermeneutik revisited. In T. Pellegrini & A. Blumauer (Hrsg.), *Semantic WebWeb. Wege zur vernetzten Wissensgesellschaft* (S. 527–533). Springer Berlin Heidelberg. Abgerufen von http://link.springer.com/chapter/10.1007/3-540-29325-6_33
- Chen, X. (2010a). Google Scholar's Dramatic Coverage Improvement Five Years after Debut. *Serials Review*, 36(4), 221–226. doi:10.1016/j.serrev.2010.08.002
- Chen, X. (2010b). The Declining Value of Subscription-based Abstracting and Indexing Services in the New Knowledge Dissemination Era. *Serials Review*, 36(2), 79–85. doi:10.1016/j.serrev.2010.02.010
- Coyle, K. (2004). Metadata: Data With a Purpose. Abgerufen 8. Dezember 2012, von http://www.kcoyle.net/meta_purpose.html
- Coyle, K. (2010). *Understanding the Semantic Web: Bibliographic Data and Metadata*. American Library Association.
- De Rosa, C., Cantrell, J., Cellentani, D., Hawk, J., Jenkins, L., & Wilson, A. (2005). *Perceptions of Libraries and Information Resources (2005)*. OCLC. Abgerufen von <http://www.oclc.org/de/de/reports/2005perceptions.htm>
- Dempsey, L., & Heery, R. (1998). Metadata: A Current View of Practice and Issues. *The Journal of Documentation*, 54/2.
- Dunsire, G., Hillmann, D. I., Phipps, J., & Coyle, K. (2011). A Reconsideration of Mapping in a Semantic World. *International Conference on Dublin Core and Metadata Applications*, 0, 26–36.
- Eberhardt, J. (2011). Was ist (bibliothekarische) Sacherschließung? In *Museum, Region, Forschung: Festschrift für Rainer Springhorn*. Hrsg. von Detlev Hellfaier und Elke Treude (Bd. 7). Detmold. Abgerufen von http://www.llb-detmold.de/fileadmin/user_upload/redaktion/dokumente/texte/2011-3_Eberhardt_Sacherschliessung.pdf

- Eckert, K. (2011). SKOS: eine Sprache für die Übertragung von Thesauri ins Semantic Web. *KIM-Technology-Watch-Report (TWR)*.
- Efthimiadis, E. N. (1996). Query expansion. *Annual Review of Information Science and Technology*, 31, 121–187.
- Endres, A., & Fellner, D. W. (2000). *Digitale Bibliotheken: Informatik-Lösungen für globale Wissensmärkte*. dpunkt.
- Europeana v1.0 Group. (2012). *Europeana Data Model Definition v5.2.3*. Abgerufen von <http://pro.europeana.eu/documents/900548/bb6b51df-ad11-4a78-8d8a-44cc41810f22>
- Felicetti, A., D’Andrea, A., & Niccolucci, F. (2012). The Open Data Semantics and the (re)use of Open Information in Cultural Heritage. Abgerufen 7. Januar 2013, von http://bpnet.eclap.eu/drupal/?q=de/home&axoid=urn:axmedis:00000:obj:464c4239-95b1-4725-9bb2-58b6f8c6b359§ion=similarByText_result&cd=4
- Fellner, D. W. (1992). *Computergrafik* (2., vollst. überarb. und erw. Aufl.). Mannheim: BI-Wiss.-Verl.
- Frohmann, B. (2009). Revisiting “what is a document?”. *Journal of Documentation*, 65(2), 291–303. doi:10.1108/00220410910937624
- Gallagher, S. (2012, Juni 7). How Google and Microsoft taught search to “understand” the Web | Ars Technica. Abgerufen 5. Dezember 2012, von <http://arstechnica.com/information-technology/2012/06/inside-the-architecture-of-googles-knowledge-graph-and-microsofts-satori/>
- Gantz, J. F. (2008, März). The Diverse and Exploding Digital Universe. An Updated Forecast of Worldwide Information Growth Through 2011. IDC-Whitepaper. Abgerufen von <http://www.emc.com/collateral/analyst-reports/diverse-exploding-digital-universe.pdf>
- Gethmann, D., & Hauser, S. (2009). *Kulturtechnik Entwerfen: Praktiken, Konzepte und Medien in Architektur und Design Science* (1., Aufl.). Transcript.
- Getty Research Institute. (2009). Introduction to Metadata: Crosswalk. Abgerufen 19. Dezember 2012, von http://www.getty.edu/research/conducting_research/standards/intrometadata/crosswalks.html
- Gradmann, S. (2009). *Signal. Information. Zeichen zu den Bedingungen des Verstehens in semantischen Netzen; Antrittsvorlesung 28. Oktober 2008, Humboldt-Universität zu Berlin, Institut für Bibliotheks- und Informati-*

- onswissenschaft (Bd. 157). Berlin: Humboldt-Univ. Abgerufen von <http://edoc.hu-berlin.de/humboldt-vl/157>
- Gradmann, S. (2012a, April 24). *From Containers to Content to Context: the Changing Role of Libraries in eScience and eScholarship*. Keynote, 10th International Bielefeld Conference, Bielefeld. Abgerufen von http://conference.ub.uni-bielefeld.de/programme/presentations/Gradmann_BC2012.pdf
- Gradmann, S. (2012b, November 7). *Modell und Interpretation: zur Erschaffung von „Welt“ in digitalen Modellen*. Gehalten auf der EVA-Berlin | WORK-SHOP 1, Berlin. Abgerufen von http://www.eva-berlin.de/index.php?article_id=73&clang=0
- Hartmann, R. (2007). *Die neue Honorarordnung für Architekten und Ingenieure*. Weka-Vlg., Augsburg.
- Hassenewert, F. (2006, April 3). Lehren des Entwerfens - Eine Untersuchung über den Diskurs des Entwerfens in Entwurfslehrbüchern der Architektur von 1945 bis 2004. *OPUS-IDN/1258*. Text.Thesis.Doctoral. Abgerufen 23. November 2012, von <http://opus.kobv.de/tuberlin/volltexte/2006/1258/>
- Havemann, S., Ullrich, T., & Fellner, D. W. (2012). The Meaning of 3D Shape and Some Techniques to Extract it. In rk T.ybury (Hrsg.), *Multimedia Information Extraction* (S. 81–97). John Wiley & Sons, Inc. Abgerufen von <http://onlinelibrary.wiley.com/doi/10.1002/9781118219546.ch5/summary>
- Heath, T., & Bizer, C. (2011). *Linked data evolving the web into a global data space*. San Rafael, Calif. (1537 Fourth Street, San Rafael, CA 94901 USA): Morgan & Claypool. Abgerufen von <http://www.morganclaypool.com/doi/abs/10.2200/S00334ED1V01Y201102WBE001>
- Heery, R. (1996). Review of metadata formats. *Program: electronic library and information systems*, 30(4), 345–373. doi:10.1108/eb047236
- Hey, A. J. G., Tansley, S., & Tolle, K. M. (2009). Jim Gray on eScience: A Transformed Scientific Method. In *The fourth paradigm: data-intensive scientific discovery*. Redmond, Wash.: Microsoft Research. Abgerufen von http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th_paradigm_book_complete_lr.pdf
- Holzer, D. (2011). BIM's Seven Deadly Sins. *International Journal of Architectural Computing vol. 9 - no. 4*, 463–480. Abgerufen von http://cumincad.scix.net/cgi-bin/works/Show?_id=ijac20119408&sort=DEFAULT&search=%2Fseries%3A%22journal%22&hits=311

- Hovestadt, V. (1998). *Informationsgebäude. Ein Integrationsmodell für Architektur und Informationstechnologien - Fraunhofer IRB - baufachinformation.de*. Univ. Karlsruhe. Abgerufen von <http://www.baufachinformation.de/dissertation/Informationsgeb%C3%A4ude/40000000000502>
- Hyvönen, E. (2012). Publishing and Using Cultural Heritage Linked Data on the Semantic Web. *Synthesis Lectures on the Semantic Web: Theory and Technology*, 2(1), 1–159. doi:10.2200/S00452ED1V01Y201210WBE003
- IFLA Study Group on the Functional Requirements for Bibliographic Records, & International Federation of Library Associations and Institutions. (1998). *Functional requirements for bibliographic records: final report*. München: K.G. Saur.
- Isaac, A. (2010, November 29). *W3C Library Linked Data_Incubator Group (LLD XG)*. Gehalten auf der SWIB 10, Köln. Abgerufen von http://swib.org/swib10/vortraege/swib10_isaac.ppt
- Kazhdan, M., Bolitho, M., & Hoppe, H. (2006). Poisson surface reconstruction. In *Proceedings of the fourth Eurographics symposium on Geometry processing*. Abgerufen von http://faculty.cs.tamu.edu/schaefer/teaching/689_Fall2006/poissonrecon.pdf
- Kent, W. (1979). Limitations of record-based information models. *ACM Trans. Database Syst.*, 4(1), 107–131. doi:10.1145/320064.320070
- Keßler, M., & Rühle, S. (2008, Mai 5). Interoperable Metadatenmodelle und Repositories. Text. Abgerufen 14. Dezember 2012, von <https://eldorado.tu-dortmund.de/handle/2003/25243>
- Klein, R. (1997). *Graphische Datenverarbeitung, in 2 Bdn., Bd.2* (aktualisierte und wesentlich erweiterte Auflage.). Oldenbourg Wissenschaftsverlag.
- Lew, M. S., Sebe, N., Djeraba, C., & Jain, R. (2006). Content-based multimedia information retrieval: State of the art and challenges. *ACM Trans. Multimedia Comput. Commun. Appl.*, 2(1), 1–19. doi:10.1145/1126004.1126005
- Migayrou, F., Mennan, Z., & Centre national d’art et de culture Georges Pompidou (Paris). (2003). *Architectures non standard*. Paris: Centre Pompidou.
- Mitchell, W. J. (1978). *Computer Aided Architectural Design*. Van Nostrand Reinhold Company.
- Mølgaard, L. L. (2009). *Context based multimedia information retrieval*. Technical University of Denmark (DTU). Abgerufen von

http://orbit.dtu.dk/fedora/objects/orbit:82854/datastreams/file_5118170/content

Naisbitt, J. (1982). *Megatrends: Ten New Directions Transforming Our Lives* (1st Aufl.). Warner Books, Inc.

Nowakowski, P., Ciepiela, E., Harężlak, D., Kocot, J., Kasztelnik, M., Bartyński, T., ... Malawski, M. (2011). The Collage Authoring Environment. *Procedia Computer Science*, 4(0), 608–617. doi:10.1016/j.procs.2011.04.064

Ostrowski, F., & Christoph, P. (2011, November). *Einführung in Linked Open Data*. SWIB 2011 Pre-Conference Workshop, Hamburg. Abgerufen von <http://swib.org/swib11/vortraege/swib11-felix-ostrowski.pdf>

Peters, I. (2009). *Folksonomies. Indexing and Retrieval in Web 2.0*. Berlin: Saur de Gruyter.

Petras, V., & Clough, P. D. (2011). Introduction to the CLEF 2011 Labs. Gehalten auf der CLEF (Notebook Papers/Labs/Workshop). Abgerufen von <http://dblp.uni-trier.de/rec/bibtex/conf/clef/PetrasC11>

Pfarr, M. (2010). *Dokumentationssystem für digitale Rekonstruktionen am Beispiel der Grabanlage Zhaoling, Provinz Shaanxi, China*. Darmstadt, Techn. Univ. Abgerufen von <http://d-nb.info/1009652060>

Rehäuser, J., & Krcmar, H. (1996). Wissensmanagement im Unternehmen. In *Schreyögg, G./Conrad, P. (Hrsg.): Managementforschung 6: Wissensmanagement* (S. 1–40). Berlin/New York: de Gruyter. Abgerufen von [http://www.krcmar.informatik.tu-muenchen.de/lehrstuhl%5Cpublikat.nsf/intern01/FC0F0EC41403EF3D412566500029C4A5/\\$FILE/96-14.pdf](http://www.krcmar.informatik.tu-muenchen.de/lehrstuhl%5Cpublikat.nsf/intern01/FC0F0EC41403EF3D412566500029C4A5/$FILE/96-14.pdf)

Reichle, I., Siegel, S., & Spelten, A. (2008). *Visuelle Modelle* (1., Aufl.). Fink (Wilhelm).

Schmitt, G. (1993). *Architectura et machina: Computer aided architectural design und virtuelle Architektur*. Vieweg & Sohn.

Schmitt, I. (2005). *Ähnlichkeitssuche in Multimedia-Datenbanken: Retrieval, Suchalgorithmen und Anfragebehandlung*. Oldenbourg Wissenschaftsverlag.

Shilane, P., Min, P., Kazhdan, M., & Funkhouser, T. (2004). The princeton shape benchmark. In *Shape Modeling Applications, 2004. Proceedings* (S. 167–178).

- Smith, M. (2009). *Final Report for the MIT FACADE Project*. Massachusetts Institute of Technology. Abgerufen von <http://facade.mit.edu/files/FACADEFinalReport.pdf>
- Stachowiak, H. (1973). *Allgemeine Modelltheorie*. Springer-Verlag.
- Steinmann, F. (2004). Modellbildung und computergestütztes Modellieren in frühen Phasen des architektonischen Entwurfs, Generalised models and computer aided modelling in early stages of architectural design. Abgerufen von <http://e-pub.uni-weimar.de/opus4/frontdoor/index/index/docId/41>
- Stock, W. G., & Stock, M. (2008). *Wissensrepräsentation: Informationen auswerten und bereitstellen*. Oldenbourg Verlag.
- Stuart, D. (2011). *Facilitating Access to the Web of Data: A Guide for Librarians*. Facet Publishing.
- Tangelder, J. W., & Veltkamp, R. C. (2008). A survey of content based 3D shape retrieval methods. *Multimedia Tools Appl.*, 39(3), 441–471. doi:10.1007/s11042-007-0181-0
- Thurrow, T. (2004). *Digitaler Architekturbestand - Untersuchungen zur computergestützten, schrittweisen Erfassung und Abbildung der Geometrie von Gebäuden im Kontext der planungsrelevanten Bauaufnahme*. Bauhaus-Universität Weimar. Abgerufen von <http://e-pub.uni-weimar.de/opus4/frontdoor/index/index/docId/140>
- Umlauf, K. (2000). Inhaltserschließung in Bibliotheken, *Berliner Handreichungen zur Bibliothekswissenschaft. Institut für Bibliothekswissenschaft der Humboldt-Universität zu Berlin*(82). Abgerufen von <http://www.ib.hu-berlin.de/~kumlau/handreichungen/h82/>
- Umlauf, K., & Gradmann, S. (2011). *Lexikon der Bibliotheks- und Informationswissenschaft* □: LBI. Stuttgart □: Hierseemann.
- Van Harmelen, F. (2006). Semantic Web Research anno 2006: main streams, popular fallacies, current status and future challenges. *Cooperative Information Agents X*, 1–7.
- Varma, M., & Zisserman, A. (2009). A Statistical Approach to Material Classification Using Image Patch Exemplars. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(11), 2032 –2047. doi:10.1109/TPAMI.2008.182
- Voß, Jakob. (2009). Zur Neubestimmung des Dokumentenbegriffs im rein Digitalen. *LIBREAS.Library Ideas*, Jg. 5, H. 2 (15). Abgerufen von <http://libreas.eu/ausgabe15/texte/002.htm>

- W3C Incubator Group. (2011, Oktober 25). Library Linked Data Incubator Group Final Report. Abgerufen 4. Dezember 2012, von <http://www.w3.org/2005/Incubator/lld/XGR-lld-20111025/>
- Wessel, R., Blümel, I., & Klein, R. (2008). The Room Connectivity Graph: Shape Retrieval in the Architectural Domain. *Proceedings of the 16-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision '2008*. Abgerufen von <http://citeseerx.ist.psu.edu/viewdoc/download?rep=rep1&type=pdf&doi=10.1.1.186.1765>
- Wessel, R., Blümel, I., & Klein, R. (2009). A 3D Shape Benchmark for Retrieval and Automatic Classification of Architectural Data. *Proceedings of Eurographics 2009 Workshop on 3D Object Retrieval*, 53–56.
- Wiesenmüller, Heidrun. (2004). Informationsaufbereitung I: Formale Erfassung. In R. Kuhlen, T. Seeger, & D. Strauch (Hrsg.), *Grundlagen der praktischen Information und Dokumentation* (S. 167–177). De Gruyter Saur.
- Wood, John. (2010). *Riding the wave - How Europe can gain from the rising tide of scientific data - Final report of the High Level Expert Group on Scientific Data* (A submission to the European Commission). Abgerufen von <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>

Anhang

Eine Vorgängerversion des Modells wurde als relationale Metadaten-Datenbank im Kontext des vorgenannten PROBADO-Projekts implementiert¹⁰². An dieser Stelle finden sich einige ausführliche Informationen zum ER-Modell, die in Kapitel 4 aus Platzgründen nicht berücksichtigt werden konnten.

Entitäten-Übersicht

Entity	Description			
Attribute	Data Type	is NULL	Description	Accepted values, examples, limitations

MODELINFO	contains primary model information				
MODELIN-FOID	int	FALSE	primary key. Unique model identification. Generated automatically		
TITLE	nvarchar	FALSE	model title		
MODELTYPEID	int	FALSE	type of model, e.g. object or building.	references catalog MODELTYPE. Currently possible values [nvarchar]: Object model Building model Environmental model Other	
MODELDESCRIPTION	nvarchar	TRUE	description of model		
LOCATION	nvar	FALSE	model location in the file	only for internal purposes	

¹⁰² Server / Datenbank: MS SQL Server 2008; Datenbankabstraktion (via Object-Relational Mapper) über Entity Framework (EF); Ansprache der Datenbank über ADO.NET Entity Framework Provider für SQL Server 2008; Weiterreichung von Suchanfragen und allgemeines Kommunikationsframework: Windows Communication Foundation (WCF); Abfragen über LINQ (Language Integrated Query). Programmiersprache: C# mit .NET.

	char		system		
IDENTIFIER_URL	nvarchar	TRUE	reference to external location	z.B. http://www.archibaseplanet.com/download/1713.html	
IDENTIFIER_DOI	nvarchar	TRUE	unambiguous reference to original file (DOI, URN)		
ACCESSIBLE	int	FALSE	indicates if and where model can be downloaded	Currently possible values: Not accessible for download Accessible for download (from internal system) Accessible for download (from external system)	
LICENSE	int	TRUE	reference to licence type (in LICENSE table)		
DATEACCEPTED	date	TRUE	date that the publisher accepted the resource into their system.		
DATEAVAILABLE	date	FALSE	date the resource is made publicly available		
DATECOPYRIGHTED	date	TRUE	specific, documented date at which the resource receives a copyrighted status, if applicable		
DATECREATED	date	TRUE	date the resource itself was put together; this could be a date range or a single date for a final component, e.g., the finalised file with all of the data		
DATEMODIFIED	date	TRUE	date of last modification		
DATEISSUED	date	TRUE	date that the resource is published or distributed e.g. to a data center		
DATEUPDATED	date	TRUE	Date of the last update to the resource, when the resource is being added to		
PREVIEW_URL	nvarchar	FALSE	html preview page in PROBADO3D	"http://3d.probado.igd.fraunhofer.de/Probado3D/ModelDetails.aspx?ID=" + MODELINFOID	
PREVIEW_JPG	nvarchar	FALSE	image preview	"http://3d.probado.igd.fraunhofer.de/Probado3D/Services/Preview.svc/Models(" + MODELINFOID + ")/thumbnail.jpg"	
PREVIEW_3DPD	nvarchar	FALSE	3d preview	"http://3d.probado.igd.fraunhofer.de/Probado3D/Services/Pre	

F	char			view.svc/Models(" + MODELIN-FOID + ")/preview.pdf"	
LICENSE					
LICENSEID	int	FALSE	primary key. Unique licence type identification. Generated automatically		
NAME	nvar char	FALSE	license name in shortened format	Currently possible values: no CC license CC BY 3.0 CC BY 3.0/de CC BY-NC 3.0 CC BY-NC 3.0/de CC BY-SA 3.0 CC BY-SA 3.0/de CC BY-NC-SA 3.0 CC BY-NC-SA 3.0/de CC BY-ND 3.0 CC BY-ND 3.0/de CC BY-NC-ND 3.0 CC BY-NC-ND 3.0/de CCO 1.0 Universal (CCO 1.0) external license	
DESCRIPTION	nvar char	TRUE	full license name with explanation	Currently possible values: keine CC-Lizenz Attribution 3.0 Namensnennung 3.0 Deutschland Attribution-NonCommercial 3.0 Unported Namensnennung-Nicht-kommerziell 3.0 Deutschland Attribution-ShareAlike 3.0 Unported Namensnennung-Weitergabe unter gleichen Bedingungen 3.0 Deutschland Attribution-NonCommercial-ShareAlike 3.0 Unported 3.0 Namensnennung-Nicht-kommerziell-Weitergabe unter gleichen Bedingungen 3.0 Deutschland Attribution-NoDerivs 3.0 Unported Namensnennung-Keine Bearbeitung 3.0 Deutschland Attribution-NonCommercial-NoDerivs 3.0 Unported	

				Namensnennung-NichtKommerziell-KeineBearbeitung 3.0 Deutschland Public Domain Dedication regard external data providers policy	
URI	nvar char	TRUE	reference to the license agreement (external URL), cc:license, xmlns:cc="http://creativecommons.org/ns#"		
RELATION	for identifying related 3D model resources				
RELATIONID	int	FALSE	primary key		
MODELIN-FOID	int	FALSE	foreign key from table MODELINFO		
RELATED-MODELIN-FOID	int	FALSE	related 3D model resource (foreign key)	3D models related to initial model according to relation defined in entity RELATIONROLE	
RELATION-ROLE			references table RELATION-ROLE (foreign key)		
RELATION-ROLE					
RELATION-ROLEID	int	FALSE	primary key	RELATIONROLEID	
RELATION-NAME	nvar char	FALSE	type of relation	References catalog relation. Allowed values: - IsSupplementTo (indicates the relation to the work // - IsSupplementedBy (indicates the relation to the work(s) which are supplements of the resource) to which the resource is a supplement) - Continues (indicates the resource is a continuation of the work referenced by the related identifier) // - IsContinuedBy (indicates the resource is continued by the work referenced by the related identifier) - IsNewVersionOf (indicates the resource is a new edition of an	

				<p>old resource, where the new edition has been modified or updated) // - IsPreviousVersionOf (indicates the resource is a previous edition of a newer resource)</p> <p>- IsPartOf (indicates the resource is a portion of another resource) // - HasPart (indicates the resource is a container of another resource)</p> <p>IsReferencedBy (indicates the resource is used as a source of information by another resource)</p> <p>- IsDocumentedBy (indicates the work is documentation about/explaining the resource referenced by the related identifier) // - Documents (indicates the relation to the work which is documentation about/explaining the resource)</p> <p>- References (indicates the relation to the work which is used as a source of information of the resource)</p> <p>- Compiles (indicates the resource is used for creating another resource or dataset) // - IsCompiledBy (indicates the resource or data is compiled/created by using another resource or dataset)</p> <p>- IsVariantFormOf (indicates the resource is a variant or different form of another resource, e.g. calculated or calibrated form or different packaging)</p> <p>- IsOriginalFormOf (indicates the relation to the works which are variant or different forms of the resource)</p>	
MODELATTRIBUTE	contains values for query engine attributes (derived from converted obj reference model) for models e.g. number of floors --> technical attributes				
MODELATTRIBUTEID	int	FALSE	primary key		

MODKEYID	int	FALSE	references table modkey (foreign key)		
MODELIN-FOID	int	FALSE	references table modelinfo (foreign key)		
ATTVALUE	nvarchar	FALSE	value of attribute		
CONTRIBUTORID	int	FALSE	references table CONTRIBUTOR (contributorid)	only for internal purposes	
MOD-KEY_LANG	tables to dynamically represent characteristic attributes of models detected by query engines e.g. number of floors, number of staircases additional table MODKEY_LANG for support of multi linguality. MODKEY is used to define the various attributes - the attribute values are stored in MODELATTRIBUTE.				
MODKEYID	int	FALSE	primary key		
MOD-KEY_LANGID	int	FALSE	characteristic attribute of models detected by query engines (derived from converted obj reference model)	references table MOD-KEY_LANG (foreign key) Example: A building has 5 floors => Define Attribute Number of floors in MOD-KEY/MODKEY_LANG and store value 5 in MODELATTRIBUTE.	
MOD-ELINFO2SUBJECTAREA	contains subjectareas related to a model (e.g. Architektur. Bauingenieurwesen...)				
SUBJECTAREAID	int	FALSE	subjectarea model belongs to	references catalog SUBJECTAREA. Currently possible values [nvarchar]: Architecture Technology / Engineering Other	
MOD-ELINFO2DISSEMINATION	for determination to which external portals or systems models are assigned				
MODELIN-FOID	int	FALSE	references table modelinfo (foreign key)		
DISSEMINATIONSYSTEMID	int	FALSE	references table allocation-system (foreign key)		
DISSEMINATION	external portals or systems (e.g. Europeana, LTP-Systems, ...)				

DISSEMINATIONID	int	FALSE	primary key		
NAME	nvarchar	FALSE	system appellation	Currently possible values: Europeana GetInfo Long-term Preservation @TIB	

MODELFILE	file information concerning models - one model can be available in various file formats				
MODEL-FILEID	int	FALSE	primary key		
MODELINFOID	int	FALSE	references table modelinfo (foreign key)		
FILENAME	nvarchar	FALSE	filename (without extension)		
FORMATID	int	FALSE	references table FORMAT (foreign key)		
SOFTWAREID	int	TRUE	references table SOFTWARE (foreign key)		
FILESIZE	int	TRUE	file size of specific model in kilobytes		
FORMAT	table contains possible PROBADO file formats				
FORMATID	int	FALSE	primary key		
FORMAT-NAME	nvarchar	TRUE	name of format		
EXTENSION	nvarchar	FALSE	file extension of format		
FORMAT-TYPEID	int	FALSE	formattype references catalog formattype		
SOFTWARE	contains information about CAD/modelling software (vendor, version)				
SOFTWAREID	int	FALSE	primary key		
CORPORATEID	int	TRUE	software company	references table corporate (foreign key)	
SOFTWARE-NAME	nvarchar	FALSE	name of software		

SOFTWAREVERSION	nvarchar	FALSE	version of software		

CONTRIBUTOR	contributor data: a contributor is either a person or a corporate				
CONTRIBUTORID	int	FALSE	primary key		
PERSONID	int	TRUE	references table person (foreign key)		
CORPORATEID	int	TRUE	references table corporate (foreign key)		
MODELINFO2CONTRIBUTOR	representation of m:n relation between modelinfo, contributor (person, corporate) and role of contributor (publisher/creator/rights holder)				
MODELINFOID	int	FALSE			
CONTRIBUTORID	int	FALSE			
CONTRIBUTORROLEID	int	FALSE	role of contributor	Currently possible values [nvarchar]: Creator Publisher/ Distributor Rightsholder Related Contributor	
PERSON	person information, foaf:Person				
PERSONID	int	FALSE	primary key		
LASTNAME	nvarchar	FALSE	last name of person		
FIRSTNAME	nvarchar	FALSE	first name of person		
BIRTHDATE	date	TRUE	birthdate of person		
EMAIL	nvarchar	TRUE	email address of person		

EMAIL_PUBLISH	boolean	FALSE	should e-mail address be published i.e. publicly visible?	Flag for EMAIL (should it be published (TRUE) or only for internal contact purposes (FALSE)). default: FALSE	
PERSON2EXTERNALINFO					
PERSONID	int	FALSE	combined primary key (PERSONID,EXTERNALINFOID)		
EXTERNALINFOID	int	FALSE	combined primary key (PERSONID,EXTERNALINFOID)		
CORPORATE	corporate data - hierarchical table to represent hierarchical organizational structures, foaf:Organization				
CORPORATEID	int	FALSE	primary key		
PARENTID	int	TRUE	primary key	parent CORPORATE (including all attributes)	
NAME	nvarchar	FALSE	name of corporate		
URL	nvarchar	TRUE	url of corporate		
ADDRESSID	int	TRUE	references address of corporate (foreign key)		
CORPORATE2EXTERNALINFO					
CORPORATEID	int	FALSE	combined primary key (CORPORATEID ,EXTERNALINFOID)		
EXTERNALINFOID	int	FALSE	combined primary key (CORPORATEID ,EXTERNALINFOID)		

EXTERNALINFO	table contains additional information about events, places, objects, persons, corporates				
EXTERNALINFOID	int	FALSE	primary key		
CORPO-	int	TRUE	references table corporate	corporate responsible for URI	

RATEID			(foreign key)		
EXTER-NALINFO	nvar char	FALSE	text: additional information about events, objects, persons, ... from external sources	e.g. URI, or text descriptions derived e.g. from DBPedia, Geonames, Personen-normdatei, ...	

ADDRESS	table contains address information				
ADDRESSID	int	FALSE	primary key		
STREET_NUMBER	int	TRUE	primary key		
ADDITION-ALINFO	nvar char	TRUE			
ZIPCODEID	int	TRUE	references ZIP of ADDRESS (foreign key)		
COUNTRYID	int	TRUE	references COUNTRY of ADDRESS (foreign key)		
TOWNID	int	TRUE	references address of corporate (foreign key)		

EVENT	contains data about events (e.g. competition, seminar, presentation)				
EVENTID	int	FALSE	primary key		
EVENTNAME	nvar char	FALSE	name of event		
EVENTDATE	date	TRUE	date of event		
EVENTDESCRIPTION	nvar char	TRUE	description of event		
PLACEID	int	TRUE	references table place (foreign key) - location of event		
MODELINFO2EVENT	representation of m:n relation between modelinfo and event				
MODELINFOID	int	FALSE	combined primary key (MODELINFOID,EVENTID)		
EVENTID	int	FALSE	combined primary key (MODELINFOID,EVENTID)		

EVENT2CONTRIBUTOR	representation of m:n relation between event and contributor (person,corporate)				
EVENTID	int	FALSE	combined primary key (EVENTID,CONTRIBUTORID)		
CONTRIBUTORID	int	FALSE	combined primary key (EVENTID,CONTRIBUTORID)		
EVENT2EXTERNALINFO					
EVENTID	int	FALSE	combined primary key (EVENTID,EXTERNALINFOID)		
EXTERNALINFOID	int	FALSE	combined primary key (EVENTID,EXTERNALINFOID)		

MODOBJECT	material thing, e.g. built architecture or object				
MOBOBJECTNAME	nvarchar	FALSE	name of real object		
OBJECTDATE	date	TRUE	date of creation		
PLACEID	int	TRUE	references table place (location) - foreign key		
MODELINFO2OBJECT	representation of m:n relation between modelinfo and modobject				
MODELINFOID	int	FALSE	combined primary key (MODELINFOID,MODOBJECTID)		
MODOBJECTID	int	FALSE	combined primary key (MODELINFOID,MODOBJECTID)		
MODOBJECT2CONTRIBUTOR	representation of m:n relation between modelobject and contributor (person,corporate)				
MODOBJECTID	int	FALSE	combined primary key (MODOBJECTID,CONTRIBUTORID)		

CONTRIBU-TORID	int	FALSE	combined primary key (MO-DOBJEC-TID,CONTRIBUTORID)	if Modelinfo has an object and object has a contributor, this is the architect/designer	
MODOB-JECT2EXTER-NALINFO					
MODOBJEC-TID	int	FALSE	combined primary key (MO-DOBJEC-TID,EXTERNALINFROID)		
EXTER-NALINFROID	int	FALSE	combined primary key (MO-DOBJEC-TID,EXTERNALINFROID)		

PLACE	customer/object information: place of object, event, ...				
PLACEID	int	FALSE	primary key		
PLACENAME	nvar char	FALSE	name of place		
LAT	float	TRUE	latitude of place		
LON	float	TRUE	longitude of place		
ADDRESSID	int	TRUE	reference to table address (foreign key)		
PLACE2EXTE-RNALINFO					
PLACEID	int	FALSE	combined primary key (PLACEID,EXTERNALINFROID)		
EXTER-NALINFROID	int	FALSE	combined primary key (PLACEID,EXTERNALINFROID)		

SUBJECT	hierarchically structured table - contains keywords and terms used to describe models/concepts. The subject (keyword/term) itself is defined in SUBJECT. To support multi linguality the additional table SUBJECT_LANG was introduced.				
SUBJECTID	int	FALSE	PK		
DEFAULT-NAME	nvar char	TRUE	Bezeichnung		

PARENTID	int	TRUE	Referenz zu parent		
MOD- ELINFO2SUB JECT	representation of m:n relation between modelinfo and subject; indicates if manual or automatic classification				
ID	int	FALSE	primary key		
MODELIN- FOID	int	FALSE			
SUBJECTID	int	FALSE	reference to SUBJECTID in SUBJECT table		
VALID	bit	FALSE	validity of entry (1..valid/0..invalid)	Flag for valid classification	
CLASSIFICA- TION	nvar char	TRUE	type of classification (e.g. manual, testrunA,..)	Currently possible values: manual machine-aided	
SUB- JECT_SCHE ME	subjectscheme like e.g. AAT (Art&Architecture Thesaurus), DDC (Dewey-Dezimalklassifikation)				
SUB- JECTSCHE- MEID	int	FALSE	primary key		
SUB- JECTSCHE- MENAME	nvar char	FALSE	name of subject scheme	Currently possible values: Getty Art and Architecture The- saurus (AAT) Dewey Decimal Classification (DDC) Gemeinsame Normdatei (GND) GBV-Basisklassifikation (BKL)	
NOTATION	notation/ classification code (in referenced subjectscheme)				
NOTA- TIONID	int	FALSE	primary key		
SUBJECTID	int	FALSE	reference to SUBJECTID in SUBJECT table		
SUB- JECTSCHE- MEID	int	FALSE	reference to SUBJECTSCHE- MEID in SUBJECT_SCHEME table		
NOTATION	nvar char	FALSE	notation/ classification code (in referenced sub- jectscheme)		

SUBJECT_LANG	addition to table SUBJECT for multi language support				
SUBJECT_LANGUAGE	int	FALSE	primary key		
SUBJECTID	int	FALSE	references subject		
SUBJECT-NAME	nvarchar	FALSE	name of subject	Zusätzliche Spalte SUBJECTNAME um die Tabelle lesbarer zu machen – die Bezeichnung die in SUBJECTNAME eingetragen wird automatisch in einen Eintrag in SUBJECT_LANG gespiegelt. Bei der Darstellung des Schlagwortbaums werden nur Einträge aus SUBJECT_LANG gezeigt. Synonyme sind weitere Bezeichnungen zum selben Subject deren Reihenfolge > 0 ist. NOTATION für DDC o.ä.	
LANGUAGE	nvarchar	FALSE	current language		
ORDER	int	FALSE	Reihenfolge		

MODELATTRIBUTES: Im Gegensatz zum beschriebenen Metadatenmodell hat die Implementation aus pragmatischen Gründen, da die Information momentan nur an der in obj konvertierten Datei gebraucht wird, die Attributwerte für Höhenkoordinate, Einheiten, Vertices, Polygone direkt am MODELINFO.

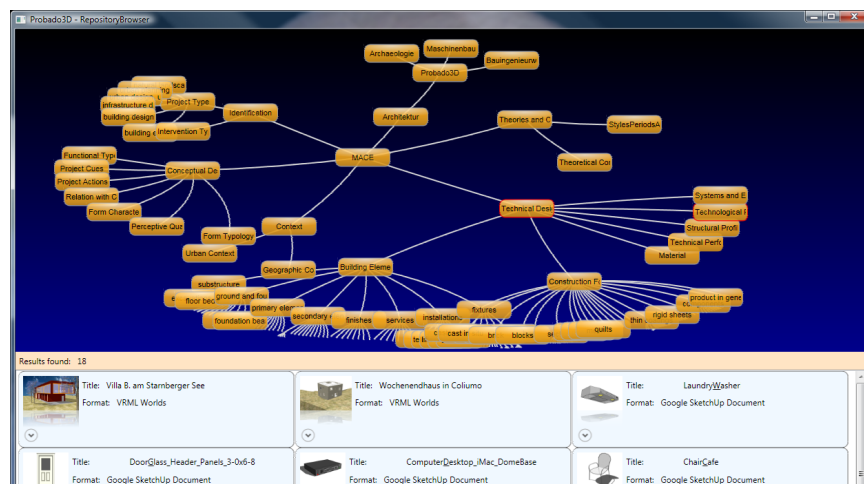


Abbildung 19: parabolisches Browsing über integrierte Klassifikation (PROBADO 3D)

SUBJECT ist in der aktuellen Implementation hierarchisch aufgebaut, da bislang nur hierarchische Klassifikationen für das in Echtzeit zu realisierendes Browsing (über einen Schlagwortbaum bzw. einen hyperbolischen Graph, vgl. Abbildung 19) integriert wurden. Es lässt Mehrsprachigkeit und Synonyme abbilden.

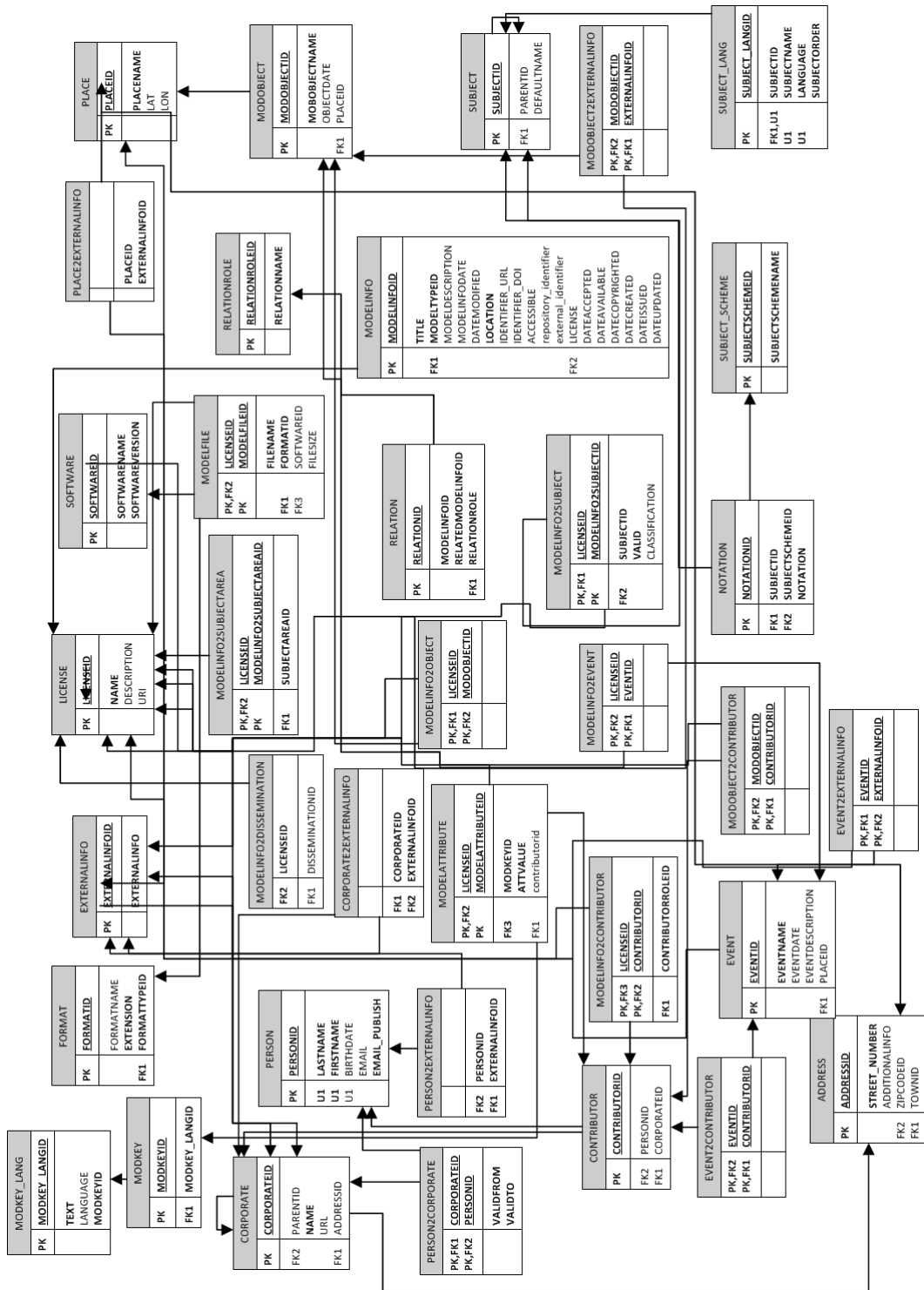


Abbildung 20: Tabellen der relationalen Datenbank

Bereitstellung/ Ausgabe

Die in der Datenbank vorgehaltenen Metadaten werden zum Zeitpunkt der Fertigstellung dieser Arbeit in folgenden Portalen verwendet bzw. in Ausgabeschnittstellen zur Verwendung bereitgestellt:

1. PROBADO 3D-Portal¹⁰³, welches zur text- und inhaltsbasierten Suche über architektonische 3D-Modelle dient und direkt an die Metadaten-Datenbank angebunden ist,
2. OAI-PMH-Schnittstelle¹⁰⁴, siehe auch XML-Ausgabe der Metadaten eines Beispielsmodells¹⁰⁵ ab der folgenden Seite. Die via OAI bereitgestellten Metadaten können von unterschiedlichen Anbietern zur Integration in eigene Portale geharvestet werden, u.a. von der Europeana, vgl. Abbildung 21.

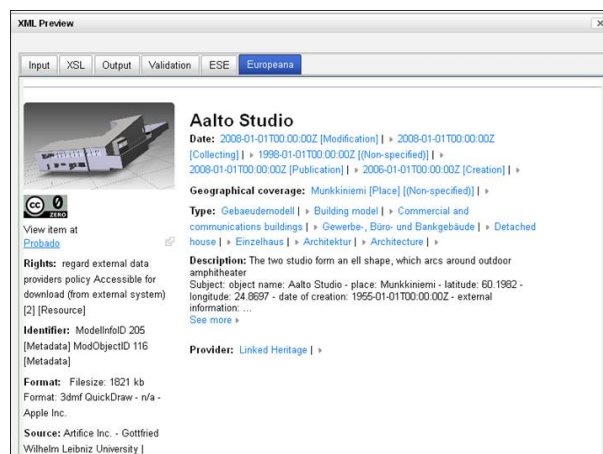


Abbildung 21: Mittels MINT-Mappingtool realisierte Vorschau auf eine Detailansicht der Modellmetadaten in der Europeana

¹⁰³ Zugang über die URL <http://www.probado.de/3d.html>

¹⁰⁴ `xsi:schemaLocation=http://www.probado.de/Probado3D/P3DM/1.0/p3dm/`
<http://www.probado.de/Probado3D/P3DM/1.0/p3dm.xsd>.

¹⁰⁵ via
<http://3d.probado.igd.fraunhofer.de/Probado3DOAI/?verb=GetRecord&metadataPrefix=p3dm&identifier=205>

```

- <OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/
    http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
  <responseDate>2013-01-03T12:26:21Z</responseDate>
  <request verb="GetRecord" identifier="205"
    metadataPrefix="p3dm">http://3d.probado.igd.fraunhofer.de/Probado3DOAI</request>
- <GetRecord>
- <record>
- <header>
  <identifier>205</identifier>
  <timestamp>2008-01-01T00:00:00Z</timestamp>
</header>
- <metadata>
- <P3DM xmlns="http://www.probado.de/Probado3D/P3DM/1.0/p3dm/"
  xsi:schemaLocation="http://www.probado.de/Probado3D/P3DM/1.0/p3dm/
    http://www.probado.de/Probado3D/P3DM/1.0/p3dm.xsd">
  <MODELINFOID>205</MODELINFOID>
  <TITLE>Aalto Studio</TITLE>
- <MODELTYPEID ID="33">
  <MODELTYPE_APPELLATION
    LANGUAGE="de">Gebaeudemodell</MODELTYPE_APPELLATION>
  <MODELTYPE_APPELLATION LANGUAGE="en">Building
    model</MODELTYPE_APPELLATION>
</MODELTYPEID>
- <SUBJECTAREAS>
- <SUBJECTAREA SUBJECTAREAID="39">
  <SUBJECTAREA_APPELLATION
    LANGUAGE="de">Architektur</SUBJECTAREA_APPELLATION>
  <SUBJECTAREA_APPELLATION
    LANGUAGE="en">Architecture</SUBJECTAREA_APPELLATION>
</SUBJECTAREA>
</SUBJECTAREAS>
  <MODELDESCRIPTION>The two studio form an ell shape, which arcs around outdoor
    amphitheater</MODELDESCRIPTION>
- <IDENTIFIERS>
  <IDENTIFIER_URL>http://www.greatbuildings.com/buildings/Aalto_Studio.html</IDENTIFIER_URL>
  <IDENTIFIER_DOI>10.5446/100000</IDENTIFIER_DOI>
</IDENTIFIERS>
  <ACCESSIBLE>Accessible for download (from external system) [2]</ACCESSIBLE>
- <LICENSE NAME="external license">
  <DESCRIPTION>regard external data providers policy</DESCRIPTION>
  <URI>n/a</URI>
</LICENSE>
- <DATES>
  <DATEMODIFIED>2008-01-01T00:00:00Z</DATEMODIFIED>
  <DATEACCEPTED>2008-01-01T00:00:00Z</DATEACCEPTED>
  <DATEUPDATED>2012-03-01T00:00:00Z</DATEUPDATED>
  <DATEAVAILABLE>2008-01-01T00:00:00Z</DATEAVAILABLE>
  <DATECOPYRIGHTED>2008-01-01T00:00:00Z</DATECOPYRIGHTED>
  <DATECREATED>2006-01-01T00:00:00Z</DATECREATED>

```

```

    <DATEISSUED>2006-01-01T00:00:00Z</DATEISSUED>
  </DATES>
- <PREVIEWS>

    <PREVIEW_URL>http://3d.probado.igd.fraunhofer.de/Probado3D/ModelDetails.aspx?
    ID=205</PREVIEW_URL>

    <PREVIEW_JPG>http://3d.probado.igd.fraunhofer.de/Probado3D/Services/Preview.svc/Models
    (205)/thumbnail.jpg</PREVIEW_JPG>

    <PREVIEW_3DPDF>http://3d.probado.igd.fraunhofer.de/Probado3D/Services/Preview.svc/Model:
    (205)/preview.pdf</PREVIEW_3DPDF>
  </PREVIEWS>
- <RELATIONS>
    <RELATION RELATIONROLE="IsSupplementTo (indicates the relation to the work to
    which the resource is a supplement)" RELATEDMODELINFOID="3145" />
  </RELATIONS>
- <MODELATTRIBUTES>
  - <MODELATTRIBUTE KEYID="4" VALUE="2358">
    <KEY_APPELLATION LANGUAGE="de">Vertizes</KEY_APPELLATION>
    <KEY_APPELLATION LANGUAGE="en">vertices</KEY_APPELLATION>
  </MODELATTRIBUTE>
  - <MODELATTRIBUTE KEYID="5" VALUE="3292">
    <KEY_APPELLATION LANGUAGE="de">Polygone</KEY_APPELLATION>
    <KEY_APPELLATION LANGUAGE="en">polygons</KEY_APPELLATION>
  </MODELATTRIBUTE>
</MODELATTRIBUTES>
- <DISSEMINATIONS>
  <string>Europeana</string>
</DISSEMINATIONS>
- <CONTRIBUTORS>
  - <CONTRIBUTOR CONTRIBUTORID="416" CONTRIBUTORROLE="Publisher/ Distributor">
    - <PERSON PERSONID="10562">
      <LASTNAME>Matthews</LASTNAME>
      <FIRSTNAME>Kevin</FIRSTNAME>
      <EMAIL>matthews@artifice.com</EMAIL>
    - <CORPORATE CORPORATEID="43">
      <NAME>Artifice Inc.</NAME>
      <URL>http://www.artifice.com</URL>
    - <ADDRESS ADDRESSID="18">
      <STREET>P.O. Box 1588</STREET>
      <ZIP>OR 97440-1588</ZIP>
      <TOWN>Eugene</TOWN>
      <COUNTRY>US</COUNTRY>
      <ADDITIONALINFO>n/a</ADDITIONALINFO>
    </ADDRESS>
    <EXTERNALINFOS />
    </CORPORATE>
    <EXTERNALINFOS />
    </PERSON>
  </CONTRIBUTOR>
</CONTRIBUTORS>

```



```

- <MODELFILES>
- <MODELFILE>
  <FILENAME>Aalto_Studio</FILENAME>
  <EXTENSION>3dmf</EXTENSION>
  <SOFTWARENAME>QuickDraw</SOFTWARENAME>
  <SOFTWAREVERSION>n/a</SOFTWAREVERSION>
  <NAME>Apple Inc.</NAME>
  <FILESIZE SIZE="1821" UNIT="kb" />
</MODELFILE>
</MODELFILES>
- <SUBJECTS>
- <SUBJECT SUBJECTID="5467">
- <NOTATIONS>
  <NOTATION VALUE="300005147" SUBJECTSCHEMENAME="Getty Art and
  Architecture Thesaurus (AAT)" />
  <NOTATION VALUE="725.2" SUBJECTSCHEMENAME="Dewey Decimal
  Classification (DDC)" />
</NOTATIONS>
<VALID>true</VALID>
<CLASSIFICATION>manual</CLASSIFICATION>
- <SUBJECT_APPELLATIONS>
  <APPELLATION LANGUAGE="en">Commercial and communications
  buildings</APPELLATION>
  <APPELLATION LANGUAGE="de">Gewerbe-, Büro- und
  Bankgebäude</APPELLATION>
</SUBJECT_APPELLATIONS>
</SUBJECT>
- <SUBJECT SUBJECTID="5449">
  <NOTATIONS />
  <VALID>true</VALID>
  <CLASSIFICATION>manual</CLASSIFICATION>
- <SUBJECT_APPELLATIONS>
  <APPELLATION LANGUAGE="en">Detached house</APPELLATION>
  <APPELLATION LANGUAGE="de">Einzelhaus</APPELLATION>
</SUBJECT_APPELLATIONS>
</SUBJECT>
</SUBJECTS>
- <EVENTS>
- <EVENT EVENTID="45">
  <EVENTNAME>GreatbuildingsModeling</EVENTNAME>
  <EVENTDATE>1998-01-01T00:00:00Z</EVENTDATE>
  <EVENTDESCRIPTION>The models range in detail from very simple formal
  massing models of just a few cuboid blocks, through basic interior/exterior
  spatial walkthrough models, up to detailed interior/exterior models complete
  with furnishings and landscaping.</EVENTDESCRIPTION>
- <PLACE PLACEID="84">
  <PLACENAME>Munkkiniemi</PLACENAME>
  <LAT>60.1982</LAT>
  <LON>24.8697</LON>
- <EXTERNALINFOS>
  - <EXTERNALINFO EXTERNALINFOID="12">

```

```

- <EXTERNALINFO EXTERNALINFOID="12">
  <TEXT>http://sws.geonames.org/645691/</TEXT>
- <CORPORATE CORPORATEID="145">
  <NAME>Geonames</NAME>
  <URL>http://www.geonames.org</URL>
  <EXTERNALINFOS />
</CORPORATE>
</EXTERNALINFO>
</EXTERNALINFOS>
- <ADDRESS ADDRESSID="17">
  <STREET>Tiilimäki 20, Munkkiniemi</STREET>
  <COUNTRY>FI</COUNTRY>
</ADDRESS>
</PLACE>
- <CONTRIBUTORS>
- <CONTRIBUTOR CONTRIBUTORID="581" CONTRIBUTORROLE="default">
  - <PERSON PERSONID="10562">
    <LASTNAME>Matthews</LASTNAME>
    <FIRSTNAME>Kevin</FIRSTNAME>
    <EMAIL>matthews@artifice.com</EMAIL>
  - <CORPORATE CORPORATEID="43">
    <NAME>Artifice Inc.</NAME>
    <URL>http://www.artifice.com</URL>
  - <ADDRESS ADDRESSID="18">
    <STREET>P.O. Box 1588</STREET>
    <ZIP>OR 97440-1588</ZIP>
    <TOWN>Eugene</TOWN>
    <COUNTRY>US</COUNTRY>
    <ADDITIONALINFO>n/a</ADDITIONALINFO>
    </ADDRESS>
    <EXTERNALINFOS />
  </CORPORATE>
  <EXTERNALINFOS />
  </PERSON>
</CONTRIBUTOR>
</CONTRIBUTORS>
- <EXTERNALINFOS>
- <EXTERNALINFO EXTERNALINFOID="14">
  <TEXT>http://www.greatbuildings.com/types/models/models.html</TEXT>
</EXTERNALINFO>
</EXTERNALINFOS>
</EVENT>
</EVENTS>
- <MODOBJECTS>
- <MODOBJECT MODOBJECTID="116">
  <MODOBJECTNAME>Aalto Studio</MODOBJECTNAME>
  <OBJECTDATE>1955-01-01T00:00:00Z</OBJECTDATE>
- <PLACE PLACEID="84">
  <PLACENAME>Munkkiniemi</PLACENAME>
  <LAT>60.1982</LAT>
  <LON>24.8697</LON>
- <EXTERNALINFOS>
  - <EXTERNALINFO EXTERNALINFOID="12">

```

```

    <TEXT>http://sws.geonames.org/645691/</TEXT>
  - <CORPORATE CORPORATEID="145">
    <NAME>Geonames</NAME>
    <URL>http://www.geonames.org</URL>
    <EXTERNALINFOS />
  </CORPORATE>
</EXTERNALINFO>
</EXTERNALINFOS>
- <ADDRESS ADDRESSID="17">
  <STREET>Tiilimäki 20, Munkkiniemi</STREET>
  <COUNTRY>FI</COUNTRY>
</ADDRESS>
</PLACE>
- <CONTRIBUTORS>
- <CONTRIBUTOR CONTRIBUTORID="420" CONTRIBUTORROLE="default">
  - <PERSON PERSONID="10396">
    <LASTNAME>Aalto</LASTNAME>
    <FIRSTNAME>Alvar_Hugo_Henrik</FIRSTNAME>
  - <EXTERNALINFOS>
    - <EXTERNALINFO EXTERNALINFOID="4">
      <TEXT>http://d-nb.info/gnd/118500023</TEXT>
    - <CORPORATE CORPORATEID="144">
      <NAME>Deutsche Nationalbibliothek</NAME>
      <URL>http://www.dnb.de</URL>
      <EXTERNALINFOS />
    </CORPORATE>
  </EXTERNALINFO>
</EXTERNALINFOS>
</PERSON>
</CONTRIBUTOR>
</CONTRIBUTORS>
- <EXTERNALINFOS>
- <EXTERNALINFO EXTERNALINFOID="3">
  <TEXT>http://en.wikiarquitectura.com/index.php/Alvar_Aalto_Studio_and_Home</TEXT>
</EXTERNALINFO>
- <EXTERNALINFO EXTERNALINFOID="13">
  <TEXT>http://www.archiplanet.org/wiki/Aalto_Studio</TEXT>
- <CORPORATE CORPORATEID="146">
  <NAME>Archiplanet</NAME>
  <URL>http://www.archiplanet.org/wiki/Archiplanet:About</URL>
  <EXTERNALINFOS />
</CORPORATE>
</EXTERNALINFO>
</EXTERNALINFOS>
</MODOBJECT>
</MODOBJECTS>
</P3DM>
</metadata>
</record>
</GetRecord>
</OAI-PMH>

```

Danksagung

Für die inhaltliche und praktische Unterstützung zum Vollenden der Dissertation bedanke ich mich im Besonderen bei Markus Bruns, Prof. Dr. Dieter Fellner, Gerd Fritsche, Prof. Dr. Stefan Gradmann, Regine Stein und Raoul Wessel.